GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 20:33:16 ; Search time 977.534 Seconds
(without alignments)
13479.231 Million cell updates/sec

Title: US-10-776-889-1
Perfect score: 2012
Sequence: 1 atcttaggcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7338684 seqs, 3274456166 residues

Total number of hits satisfying chosen parameters: 14677368

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	2012	100.0	2012	19	US-10-776-889-1	Sequence 1, Appli
→	2	2012	100.0	2012	19	US-10-776-311-3	Sequence 3, Appli
	3	2012	100.0	2012	22	US-10-985-109-81	Sequence 81, Appl
	4	2012	100.0	2012	22	US-10-985-254-81	Sequence 81, Appl
c	5	1961.4	97.5	12456	22	US-10-985-109-95	Sequence 95, Appl
c	6	1961.4	97.5	12456	22	US-10-985-254-95	Sequence 95, Appl
	7	1883	93.6	1883	19	US-10-776-889-13	Sequence 13, Appl
	8	1719	85.4	1719	19	US-10-776-889-14	Sequence 14, Appl
	9	1553	77.2	1553	19	US-10-776-889-15	Sequence 15, Appl
	10	1367	67.9	1367	19	US-10-776-889-16	Sequence 16, Appl
	11	1160	57.7	1160	19	US-10-776-889-17	Sequence 17, Appl
	12	967	48.1	967	19	US-10-776-889-18	Sequence 18, Appl
	13	770	38.3	770	19	US-10-776-889-19	Sequence 19, Appl
	14	584	29.0	584	19	US-10-776-889-20	Sequence 20, Appl
	15	425	21.1	425	19	US-10-776-889-21	Sequence 21, Appl
	16	193.4	9.6	1499	18	US-10-424-599-30062	Sequence 30062, A
	17	174	8.6	174	19	US-10-776-889-22	Sequence 22, Appl
	18	170	8.4	3673778	16	US-10-312-841-1	Sequence 1, Appli
	19	156	7.8	15548	15	US-10-311-455-2128	Sequence 2128, Ap
c	20	150.2	7.5	8056	20	US-10-473-126-386	Sequence 386, App
	21	148.6	7.4	8056	20	US-10-473-126-386	Sequence 386, App
	22	147.4	7.3	7597	15	US-10-311-455-986	Sequence 986, App
c	23	142.2	7.1	8056	20	US-10-473-126-240	Sequence 240, App
	24	136	6.8	115218	21	US-10-278-698-255	Sequence 255, App
	25	136	6.8	115218	21	US-10-278-698-769	Sequence 769, App
	26	133.6	6.6	8056	20	US-10-473-126-240	Sequence 240, App
c	27	133	6.6	158001	17	US-10-211-179-11	GENERAL INFORMATI
	28	131.8	6.6	7814	18	US-10-221-714A-252	Sequence 252, App
	29	131.2	6.5	43527	22	US-10-756-149-2369	Sequence 2369, Ap
	30	128.4	6.4	7498	15	US-10-311-455-230	Sequence 230, App
	31	124.4	6.2	6175	15	US-10-311-455-1280	Sequence 1280, Ap
	32	123.8	6.2	6419	15	US-10-311-455-240	Sequence 240, App
	33	123.8	6.2	12237	15	US-10-311-455-2331	Sequence 2331, Ap
c	34	122.4	6.1	1300	19	US-10-668-749A-1	Sequence 1, Appli
c	35	121.6	6.0	3673778	16	US-10-312-841-1	Sequence 1, Appli
	36	121.4	6.0	3673778	16	US-10-312-841-2	Sequence 2, Appli
c	37	120.4	6.0	3252	13	US-10-027-632-113786	Sequence 113786,
c	38	120.4	6.0	3252	13	US-10-027-632-113787	Sequence 113787,
c	39	120.4	6.0	3252	13	US-10-027-632-113788	Sequence 113788,
c	40	120.4	6.0	3252	17	US-10-027-632-113786	Sequence 113786,
c	41	120.4	6.0	3252	17	US-10-027-632-113787	Sequence 113787,
c	42	120.4	6.0	3252	17	US-10-027-632-113788	Sequence 113788,
	43	120	6.0	6109	15	US-10-311-455-299	Sequence 299, App
	44	120	6.0	6109	17	US-10-221-613-33	Sequence 33, Appl
c	45	119	5.9	74665	20	US-10-719-993-6854	Sequence 6854, Ap

ALIGNMENTS

RESULT 1
US-10-776-889-1
; Sequence 1, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BB1531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 2012
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-1

Query Match 100.0%; Score 2012; DB 19; Length 2012;
Best Local Similarity 100.0%; Pred. No. 2e-242;
Matches 2012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60

Qy         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120
```

```
Qy       121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180
Db       121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180

Qy       181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240
Db       181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240

Qy       241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300
Db       241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300

Qy       301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360
Db       301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360

Qy       361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420
Db       361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420

Qy       421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480
Db       421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480

Qy       481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540
Db       481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540

Qy       541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600
Db       541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600

Qy       601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660
Db       601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660

Qy       661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720
Db       661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720

Qy       721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780
Db       721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780

Qy       781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840
Db       781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840

Qy       841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900
Db       841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900

Qy       901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960
Db       901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960

Qy       961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020
Db       961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020

Qy      1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080
Db      1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080

Qy      1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140
Db      1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140

Qy      1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200
Db      1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200

Qy      1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260
Db      1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260

Qy      1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320
Db      1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320

Qy      1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380
Db      1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380

Qy      1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440
Db      1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440

Qy      1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500
Db      1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500

Qy      1501 GAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560
Db      1501 GAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560

Qy      1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620
Db      1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620

Qy      1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680
Db      1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680

Qy      1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740
Db      1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740

Qy      1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800
Db      1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800

Qy      1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860
Db      1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860

Qy      1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
Db      1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920

Qy      1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
Db      1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980

Qy      1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
Db      1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
```

RESULT 2
US-10-776-311-3
; Sequence 3, Application US/10776311
; Publication No. US20040172682A1
; GENERAL INFORMATION:
; APPLICANT: Edgar B. Cahoon
; APPLICANT: Howard G. Damude
; APPLICANT: William D. Hitz
; APPLICANT: Anthony J. Kinney
; APPLICANT: Charles W. Kolar
; APPLICANT: Zhan Bin Liu
; TITLE OF INVENTION: Production of Long Chain Polyunsaturated Fatty Acids in Plants
; FILE REFERENCE: BB1538 US NA
; CURRENT APPLICATION NUMBER: US/10/776,311
; CURRENT FILING DATE: 2004-02-11
; PRIOR APPLICATION NUMBER: US 60/446,941

```
;  PRIOR FILING DATE: 2003-02-12
;  NUMBER OF SEQ ID NOS: 98
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
;   LENGTH: 2012
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-776-311-3

  Query Match           100.0%;  Score 2012;  DB 19;  Length 2012;
  Best Local Similarity  100.0%;  Pred. No. 2e-242;
  Matches 2012;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60

Qy         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120

Qy        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180

Qy        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240

Qy        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300

Qy        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360

Qy        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420

Qy        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480

Qy        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540

Qy        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600

Qy        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660

Qy        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720

Qy        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780

Qy        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840

Qy        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900

Qy        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960

Qy        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020

Qy       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080

Qy       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140

Qy       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200

Qy       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260

Qy       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320

Qy       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380

Qy       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440

Qy       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500

Qy       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560

Qy       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620

Qy       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680

Qy       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740

Qy       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800

Qy       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860

Qy       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920

Qy       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980

Qy       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
```

```
                ||||||||||||||||||||||||||||||||
Db       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012


RESULT 3
US-10-985-109-81
; Sequence 81, Application US/10985109
; Publication No. US20050132441A1
; GENERAL INFORMATION:
;  APPLICANT: Yadav, Narendra
;  APPLICANT:  Damude, Howard
;  TITLE OF INVENTION: DELTA-15 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURA
;  TITLE OF INVENTION:  FATTY ACIDS IN OILSEED PLANTS AND OLEAGINOUS YEAST
;  FILE REFERENCE: CL2432
;  CURRENT APPLICATION NUMBER: US/10/985,109
;  CURRENT FILING DATE:  2004-11-10
;  PRIOR APPLICATION NUMBER: US 60/519191
;  PRIOR FILING DATE: 2003-11-12
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81
;   LENGTH: 2012
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-985-109-81

  Query Match             100.0%;  Score 2012;  DB 22;  Length 2012;
  Best Local Similarity   100.0%;  Pred. No. 2e-242;
  Matches 2012;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60

Qy         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120

Qy        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180

Qy        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240

Qy        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300

Qy        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360

Qy        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420

Qy        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480

Qy        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540

Qy        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600

Qy        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660

Qy        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720

Qy        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780

Qy        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840

Qy        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900

Qy        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960

Qy        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020

Qy       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080

Qy       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140

Qy       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200

Qy       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260

Qy       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320

Qy       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380

Qy       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440

Qy       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500

Qy       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560

Qy       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620

Qy       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680

Qy       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740
```

```
Qy       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800

Qy       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860

Qy       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920

Qy       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980

Qy       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||
Db       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012


RESULT 4
US-10-985-254-81
; Sequence 81, Application US/10985254
; Publication No. US20050132442A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. duPont de Nemours and Co., Inc.
;  APPLICANT:  Yadav, Narendra
;  APPLICANT:  Damude, Howard
;  TITLE OF INVENTION: DELTA 15 DESATURASES SUITABLE FOR ALTERING LEVELS OF
;  TITLE OF INVENTION:  POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST
;  FILE REFERENCE: CL2432
;  CURRENT APPLICATION NUMBER: US/10/985,254
;  CURRENT FILING DATE:  2004-11-10
;  PRIOR APPLICATION NUMBER: US 60/519191
;  PRIOR FILING DATE: 2003-11-12
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 81
;   LENGTH: 2012
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-985-254-81

  Query Match           100.0%;  Score 2012;  DB 22;  Length 2012;
  Best Local Similarity  100.0%;  Pred. No. 2e-242;
  Matches 2012;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60

Qy         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120

Qy        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180

Qy        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240

Qy        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300

Qy        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360

Qy        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420

Qy        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480

Qy        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540

Qy        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600

Qy        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660

Qy        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720

Qy        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780

Qy        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840

Qy        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900

Qy        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960

Qy        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020

Qy       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080

Qy       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140

Qy       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200

Qy       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260

Qy       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320

Qy       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380

Qy       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440

Qy       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500

Qy      1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560

Qy      1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620

Qy      1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680

Qy      1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740

Qy      1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800

Qy      1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860

Qy      1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920

Qy      1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980

Qy      1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||
Db      1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012


RESULT 5
US-10-985-109-95/c
; Sequence 95, Application US/10985109
; Publication No. US20050132441A1
; GENERAL INFORMATION:
;  APPLICANT: Yadav, Narendra
;  APPLICANT:  Damude, Howard
;  TITLE OF INVENTION: DELTA-15 DESATURASES SUITABLE FOR ALTERING LEVELS OF POLYUNSATURA
;  TITLE OF INVENTION:  FATTY ACIDS IN OILSEED PLANTS AND OLEAGINOUS YEAST
;  FILE REFERENCE: CL2432
;  CURRENT APPLICATION NUMBER: US/10/985,109
;  CURRENT FILING DATE:  2004-11-10
;  PRIOR APPLICATION NUMBER: US 60/519191
;  PRIOR FILING DATE: 2003-11-12
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 95
;   LENGTH: 12456
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: plasmid pKR585
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1201)..(1201)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-985-109-95

  Query Match             97.5%;  Score 1961.4;  DB 22;  Length 12456;
  Best Local Similarity   99.8%;  Pred. No. 6.5e-236;
  Matches 1974;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy        38 CACATGCAAGTTTTTATTTCAATCCCTTTTCCTTTGAATAACTGACCAAGAACAACAAG- 96
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |

Db     12431 CACATGCAAGTTTTTATTTCAATCCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGA 12372

Qy        97 -AAAAAAAAAAAAAAGAAAAGGATCATTTTGAAAGGATATTTTTCGCTCCTATTCAAATAC 155
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12371 AAAAAAAAAAAAAAAGAAAAGGATCATTTTGAAAGGATATTTTTCGCTCCTATTCAAATAC 12312

Qy       156 TGTATTTTTACCAAAAAAACTGTATTTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCG 215
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12311 TGTATTTTTACCAAAAAAACTGTATTTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCG 12252

Qy       216 ACTCTCATGATTTCCTTCATATGCCAATCACTCTATTTATAAATGGCATAAGGTAGTGTG 275
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12251 ACTCTCATGATTTCCTTCATATGCCAATCACTCTATTTATAAATGGCATAAGGTAGTGTG 12192

Qy       276 AACAATTGCAAAGCTTGTCATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCT 335
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12191 AACAATTGCAAAGCTTGTCATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCT 12132

Qy       336 TTCAAAATTATCTGCACCCCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTT 395
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12131 TTCAAAATTATCTGCACCCCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTT 12072

Qy       396 TTTCGAATAGTCATGCAGTGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACA 455
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12071 TTTCGAATAGTCATGCAGTGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACA 12012

Qy       456 CTGGCCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTA 515
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     12011 CTGGCCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTA 11952

Qy       516 AAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATT 575
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11951 AAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATT 11892

Qy       576 CTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCT 635
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11891 CTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCT 11832

Qy       636 ATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTT 695
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11831 ATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTT 11772

Qy       696 AAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTT 755
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11771 AAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTT 11712

Qy       756 AATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATA 815
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11711 AATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATA 11652

Qy       816 TGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGA 875
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11651 TGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGA 11592

Qy       876 CAGTGAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAAC 935
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11591 CAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAAC 11532

Qy       936 ATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTT 995
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11531 ATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTT 11472

Qy       996 ATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTAT 1055
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11471 ATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTAT 11412

Qy      1056 CCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACT 1115
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11411 CCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACT 11352

Qy      1116 CTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCT 1175
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     11351 CTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCT 11292
```

```
Qy       1176 CCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGT 1235
             ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db      11291 CCTCAAGTGATTTCCAACTCAGCATTGGCATCTAATTGATTCAGTATATCTATTGCATGT 11232

Qy       1236 GTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAAT 1295
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11231 GTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAAT 11172

Qy       1296 ATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATA 1355
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11171 ATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATA 11112

Qy       1356 AATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATAT 1415
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11111 AATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATAT 11052

Qy       1416 GATATTTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAAT 1475
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11051 GATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAAT 10992

Qy       1476 AAATTTTAAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGAC 1535
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10991 AAATTTTAAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGAC 10932

Qy       1536 CAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAA 1595
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10931 CAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAA 10872

Qy       1596 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAA 1655
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10871 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAA 10812

Qy       1656 CAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAA 1715
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10811 CAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAA 10752

Qy       1716 ACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGA 1775
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10751 ACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGA 10692

Qy       1776 AATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAG 1835
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10691 AATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAG 10632

Qy       1836 AGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATA 1895
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10631 AGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATA 10572

Qy       1896 ATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCA 1955
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10571 ATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCA 10512

Qy       1956 AAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10511 AAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 10455
```

```
RESULT 6
US-10-985-254-95/c
; Sequence 95, Application US/10985254
; Publication No. US20050132442A1
; GENERAL INFORMATION:
;  APPLICANT: E.I. duPont de Nemours and Co., Inc.
;  APPLICANT:  Yadav, Narendra
;  APPLICANT:  Damude, Howard
;  TITLE OF INVENTION: DELTA 15 DESATURASES SUITABLE FOR ALTERING LEVELS OF
;  TITLE OF INVENTION:  POLYUNSATURATED FATTY ACIDS IN OLEAGINOUS YEAST
;  FILE REFERENCE: CL2432
;  CURRENT APPLICATION NUMBER: US/10/985,254
;  CURRENT FILING DATE:  2004-11-10
;  PRIOR APPLICATION NUMBER: US 60/519191
;  PRIOR FILING DATE: 2003-11-12
;  NUMBER OF SEQ ID NOS: 109
;  SOFTWARE: PatentIn version 3.3
; SEQ ID NO 95
;   LENGTH: 12456
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: plasmid pKR585
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1201)..(1201)
;   OTHER INFORMATION: n is a, c, g, or t
US-10-985-254-95

  Query Match             97.5%;  Score 1961.4;  DB 22;  Length 12456;
  Best Local Similarity   99.8%;  Pred. No. 6.5e-236;
  Matches 1974;  Conservative    0;  Mismatches    1;  Indels    2;  Gaps    1;

Qy         38 CACATGCAAGTTTTTATTTCAATCCCTTTTCCTTTGAATAACTGACCAAGAACAACAAG- 96
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db      12431 CACATGCAAGTTTTTATTTCAATCCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGA 12372

Qy         97 -AAAAAAAAAAAAAAGAAAAGGATCATTTTGAAAGGATATTTTTCGCTCCTATTCAAATAC 155
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12371 AAAAAAAAAAAAAAGAAAAGGATCATTTTGAAAGGATATTTTTCGCTCCTATTCAAATAC 12312

Qy        156 TGTATTTTTACCAAAAAAACTGTATTTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCG 215
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12311 TGTATTTTTACCAAAAAAACTGTATTTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCG 12252

Qy        216 ACTCTCATGATTTCCTTCATATGCCAATCACTCTATTTATAAATGGCATAAGGTAGTGTG 275
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12251 ACTCTCATGATTTCCTTCATATGCCAATCACTCTATTTATAAATGGCATAAGGTAGTGTG 12192

Qy        276 AACAATTGCAAAGCTTGTCATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCT 335
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12191 AACAATTGCAAAGCTTGTCATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCT 12132

Qy        336 TTCAAAATTATCTGCACCCCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTT 395
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12131 TTCAAAATTATCTGCACCCCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTT 12072

Qy        396 TTTCGAATAGTCATGCAGTGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACA 455
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12071 TTTCGAATAGTCATGCAGTGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACA 12012

Qy        456 CTGGCCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTA 515
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12011 CTGGCCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTA 11952

Qy        516 AAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATT 575
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11951 AAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATT 11892

Qy        576 CTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCT 635
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11891 CTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCT 11832

Qy        636 ATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTT 695
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11831 ATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTT 11772

Qy        696 AAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTT 755
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11771 AAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTT 11712

Qy        756 AATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATA 815
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11711 AATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATA 11652

Qy        816 TGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGA 875
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11651 TGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGA 11592
```

```
Qy        876 CAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAAC 935
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11591 CAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAAC 11532

Qy        936 ATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTT 995
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11531 ATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTT 11472

Qy        996 ATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTAT 1055
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11471 ATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTAT 11412

Qy       1056 CCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACT 1115
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11411 CCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACT 11352

Qy       1116 CTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCT 1175
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11351 CTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCT 11292

Qy       1176 CCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGT 1235
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db      11291 CCTCAAGTGATTTCCAACTCAGCATTGGCATCTAATTGATTCAGTATATCTATTGCATGT 11232

Qy       1236 GTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAAT 1295
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11231 GTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAAT 11172

Qy       1296 ATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATA 1355
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11171 ATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATA 11112

Qy       1356 AATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATAT 1415
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11111 AATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATAT 11052

Qy       1416 GATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAAT 1475
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11051 GATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAAT 10992

Qy       1476 AAATTTTAAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGAC 1535
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10991 AAATTTTAAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGAC 10932

Qy       1536 CAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAA 1595
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10931 CAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAA 10872

Qy       1596 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAA 1655
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10871 CAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAA 10812

Qy       1656 CAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAA 1715
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10811 CAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAA 10752

Qy       1716 ACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGA 1775
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10751 ACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGA 10692

Qy       1776 AATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAG 1835
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10691 AATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAG 10632

Qy       1836 AGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATA 1895
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10631 AGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATA 10572

Qy       1896 ATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCA 1955
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10571 ATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCA 10512
```

```
Qy       1956 AAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      10511 AAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 10455
```

RESULT 7
US-10-776-889-13
; Sequence 13, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BB1531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1883
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-13

Query Match 93.6%; Score 1883; DB 19; Length 1883;
Best Local Similarity 100.0%; Pred. No. 2.6e-226;
Matches 1883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        130 GATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAAACTGTATTTTTCCTAC 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAAACTGTATTTTTCCTAC 60

Qy        190 ACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCCAATCACTCT 249
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCCAATCACTCT 120

Qy        250 ATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAAAGCTTGCAA 309
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAAAGCTTGCAA 180

Qy        310 TGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGCTATTAATCT 369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGCTATTAATCT 240

Qy        370 AACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTAATTTCCCCG 429
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTAATTTCCCCG 300

Qy        430 TGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCATTACATGAAA 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCATTACATGAAA 360

Qy        490 AGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTA 549
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTA 420

Qy        550 ATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCA 609
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCA 480

Qy        610 TATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAG 669
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAG 540

Qy        670 ACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTT 729
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTT 600

Qy        730 CTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTTGTAAATTTTTATTTATC 789
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTTGTAAATTTTTATTTATC 660

Qy        790 AATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATC 849
```

```
Db        661 AATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATC 720
Qy        850 ATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTT 909
Db        721 ATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTT 780
Qy        910 ATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATT 969
Db        781 ATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATT 840
Qy        970 TCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTT 1029
Db        841 TCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTT 900
Qy       1030 TTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAA 1089
Db        901 TTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAA 960
Qy       1090 AATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTG 1149
Db        961 AATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTG 1020
Qy       1150 AAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTG 1209
Db       1021 AAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTG 1080
Qy       1210 ATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATT 1269
Db       1081 ATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATT 1140
Qy       1270 AATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTAT 1329
Db       1141 AATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTAT 1200
Qy       1330 TTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGA 1389
Db       1201 TTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGA 1260
Qy       1390 TTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTT 1449
Db       1261 TTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTT 1320
Qy       1450 ATAAGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAA 1509
Db       1321 ATAAGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAA 1380
Qy       1510 AAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATT 1569
Db       1381 AAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATT 1440
Qy       1570 CCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCAT 1629
Db       1441 CCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCAT 1500
Qy       1630 CTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATC 1689
Db       1501 CTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATC 1560
Qy       1690 ACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCC 1749
Db       1561 ACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCC 1620
Qy       1750 GTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAA 1809
Db       1621 GTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAA 1680
Qy       1810 CGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACA 1869
Db       1681 CGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACA 1740
Qy       1870 TGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATT 1929
Db       1741 TGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATT 1800
Qy       1930 TCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAG 1989
Db       1801 TCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAG 1860
Qy       1990 GTTAACTAAGAAGCAATACTTCA 2012
Db       1861 GTTAACTAAGAAGCAATACTTCA 1883
```

RESULT 8
US-10-776-889-14
; Sequence 14, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BB1531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-14

Query Match 85.4%; Score 1719; DB 19; Length 1719;
Best Local Similarity 100.0%; Pred. No. 8.1e-206;
Matches 1719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        294 CATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACC 353
Db          1 CATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACC 60
Qy        354 CCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAG 413
Db         61 CCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAG 120
Qy        414 TGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGT 473
Db        121 TGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGT 180
Qy        474 CCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTA 533
Db        181 CCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTA 240
Qy        534 AATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCA 593
Db        241 AATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCA 300
Qy        594 AATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCT 653
Db        301 AATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCT 360
Qy        654 AAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTAT 713
Db        361 AAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTAT 420
Qy        714 ATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTG 773
Db        421 ATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTG 480
Qy        774 TAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTG 833
Db        481 TAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTG 540
Qy        834 TTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGAT 893
Db        541 TTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGAT 600
```

```
Qy        894 GAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATT 660

Qy        954 TCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCAT 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCAT 720

Qy       1014 TTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACT 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACT 780

Qy       1074 CATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAAT 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAAT 840

Qy       1134 CATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAAC 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAAC 900

Qy       1194 TCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAA 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAA 960

Qy       1254 TACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAA 1313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAA 1020

Qy       1314 AATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAA 1373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1021 AATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAA 1080

Qy       1374 TGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGG 1433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1081 TGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGG 1140

Qy       1434 TTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTA 1493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1141 TTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTA 1200

Qy       1494 TATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCT 1553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1201 TATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCT 1260

Qy       1554 TAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGA 1613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1261 TAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGA 1320

Qy       1614 AGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAAC 1673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1321 AGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAAC 1380

Qy       1674 CACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGA 1733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1381 CACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGA 1440

Qy       1734 AGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTT 1793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1441 AGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTT 1500

Qy       1794 TCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGT 1853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1501 TCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGT 1560

Qy       1854 ACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAA 1913
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1561 ACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAA 1620

Qy       1914 CTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCT 1973
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1621 CTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCT 1680

Qy       1974 GAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
              |||||||||||||||||||||||||||||||||||||||
Db        1681 GAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 1719
```

RESULT 9
US-10-776-889-15
; Sequence 15, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BB1531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1553
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-15

Query Match 77.2%; Score 1553; DB 19; Length 1553;
Best Local Similarity 100.0%; Pred. No. 4.4e-185;
Matches 1553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        460 CCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAA 519
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAA 60

Qy        520 TATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTAT 579
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTAT 120

Qy        580 TAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATAT 639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATAT 180

Qy        640 ACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAG 699
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAG 240

Qy        700 CTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATA 759
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATA 300

Qy        760 AGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTT 819
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTT 360

Qy        820 TTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGT 879
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGT 420

Qy        880 GAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGT 939
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGT 480

Qy        940 ATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTT 999
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTT 540

Qy       1000 TTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCT 1059
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCT 600

Qy       1060 CTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTAT 1119
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTAT 660

Qy       1120 CTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTC 1179
```

```
Db        661 CTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTC 720

Qy       1180 AAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAA 1239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAA 780

Qy       1240 AAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTT 1299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTT 840

Qy       1300 TTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATC 1359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATC 900

Qy       1360 TAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATA 1419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATA 960

Qy       1420 TTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAAT 1479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAAT 1020

Qy       1480 TTTAAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGC 1539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTTAAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGC 1080

Qy       1540 ATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGA 1599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGA 1140

Qy       1600 TAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGA 1659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGA 1200

Qy       1660 AAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAA 1719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAA 1260

Qy       1720 GTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATA 1779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATA 1320

Qy       1780 GTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTG 1839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTG 1380

Qy       1840 CACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCC 1899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCC 1440

Qy       1900 TGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAAT 1959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAAT 1500

Qy       1960 GAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 1553
```

RESULT 10
US-10-776-889-16
; Sequence 16, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: Kinney, Anthony
; TITLE OF INVENTION: Seed Specific Promoters
; FILE REFERENCE: BB1531 US NA
; CURRENT APPLICATION NUMBER: US/10/776,889
; CURRENT FILING DATE: 2004-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Office 97

; SEQ ID NO 16
; LENGTH: 1367
; TYPE: DNA
; ORGANISM: Glycine max
US-10-776-889-16

Query Match 67.9%; Score 1367; DB 19; Length 1367;
Best Local Similarity 100.0%; Pred. No. 7.5e-162;
Matches 1367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        646 ATATTTCTAAATTATTAAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATAT 705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATATTTCTAAATTATTAAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATAT 60

Qy        706 CATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAAT 765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAAT 120

Qy        766 AAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCA 825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCA 180

Qy        826 CATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAA 885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAA 240

Qy        886 TAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAAACCTAAACAAACATGTATCTAC 945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAAACCTAAACAAACATGTATCTAC 300

Qy        946 TCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTAT 1005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTAT 360

Qy       1006 CATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTA 1065
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTA 420

Qy       1066 TTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTC 1125
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTC 480

Qy       1126 CACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGA 1185
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGA 540

Qy       1186 TTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCT 1245
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCT 600

Qy       1246 TTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTT 1305
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTT 660

Qy       1306 CTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATA 1365
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATA 720

Qy       1366 TATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTT 1425
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTT 780

Qy       1426 TATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAA 1485
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAA 840

Qy       1486 TATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTT 1545
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTT 900

Qy       1546 TACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCG 1605
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCG 960
```

```
Qy       1606 TTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGA 1665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGA 1020

Qy       1666 CGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCA 1725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCA 1080

Qy       1726 ACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTA 1785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTA 1140

Qy       1786 TTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATT 1845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATT 1200

Qy       1846 TGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAA 1905
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAA 1260

Qy       1906 GGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTC 1965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTC 1320

Qy       1966 ACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 1367


RESULT 11
US-10-776-889-17
; Sequence 17, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Zhan-Bin
;  APPLICANT:  Kinney, Anthony
;  TITLE OF INVENTION: Seed Specific Promoters
;  FILE REFERENCE: BB1531 US NA
;  CURRENT APPLICATION NUMBER: US/10/776,889
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
;   LENGTH: 1160
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-776-889-17

  Query Match             57.7%;  Score 1160;  DB 19;  Length 1160;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-136;
  Matches 1160;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        853 ATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATT 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATT 60

Qy        913 TCCTAAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCTAAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCA 120

Qy        973 TTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTA 180

Qy       1033 CTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAAT 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAAT 240

Qy       1093 TGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAA 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAA 300

Qy       1153 AGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATT 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATT 360

Qy       1213 GATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAAT 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAAT 420

Qy       1273 CTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTT 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTT 480

Qy       1333 TTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTA 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTA 540

Qy       1393 AACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATA 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATA 600

Qy       1453 AGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAA 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAA 660

Qy       1513 TATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCT 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCT 720

Qy       1573 ATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTC 1632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTC 780

Qy       1633 TTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACT 1692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACT 840

Qy       1693 CTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTT 1752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTT 900

Qy       1753 TCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGT 1812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGT 960

Qy       1813 GCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGA 1872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGA 1020

Qy       1873 AAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCA 1932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCA 1080

Qy       1933 CTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTT 1992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTT 1140

Qy       1993 AACTAAGAAGCAATACTTCA 2012
              ||||||||||||||||||||
Db       1141 AACTAAGAAGCAATACTTCA 1160


RESULT 12
US-10-776-889-18
; Sequence 18, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Zhan-Bin
;  APPLICANT:  Kinney, Anthony
;  TITLE OF INVENTION: Seed Specific Promoters
;  FILE REFERENCE: BB1531 US NA
;  CURRENT APPLICATION NUMBER: US/10/776,889
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 22
```

```
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
;   LENGTH: 967
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-776-889-18

  Query Match             48.1%;  Score 967;  DB 19;  Length 967;
  Best Local Similarity   100.0%;  Pred. No. 6.8e-112;
  Matches  967;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1046 TTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTT 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTT 60

Qy       1106 TATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGT 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGT 120

Qy       1166 CTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATC 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATC 180

Qy       1226 TATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAA 1285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAA 240

Qy       1286 AGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTA 1345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTA 300

Qy       1346 TATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAA 1405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAA 360

Qy       1406 TATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATA 1465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATA 420

Qy       1466 TGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATA 1525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATA 480

Qy       1526 AATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTG 1585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTG 540

Qy       1586 AAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTT 1645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTT 600

Qy       1646 TTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTT 1705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTT 660

Qy       1706 TTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACA 1765
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACA 720

Qy       1766 CTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCT 1825
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCT 780

Qy       1826 TATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTC 1885
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTC 840

Qy       1886 TAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGT 1945
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGT 900

Qy       1946 GACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAA 2005
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        901 GACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAA 960

Qy       2006 TACTTCA 2012
              |||||||
Db        961 TACTTCA 967


RESULT 13
US-10-776-889-19
; Sequence 19, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Zhan-Bin
;  APPLICANT:  Kinney, Anthony
;  TITLE OF INVENTION: Seed Specific Promoters
;  FILE REFERENCE: BB1531 US NA
;  CURRENT APPLICATION NUMBER: US/10/776,889
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 19
;   LENGTH: 770
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-776-889-19

  Query Match             38.3%;  Score 770;  DB 19;  Length 770;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-87;
  Matches  770;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1243 TCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTT 1302
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTT 60

Qy       1303 TTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAA 1362
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAA 120

Qy       1363 ATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTT 1422
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTT 180

Qy       1423 TTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTT 1482
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTT 240

Qy       1483 AAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATA 1542
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATA 300

Qy       1543 TTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAA 1602
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAA 360

Qy       1603 TCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAG 1662
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAG 420

Qy       1663 GGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTC 1722
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTC 480

Qy       1723 TCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTA 1782
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTA 540

Qy       1783 GTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCAC 1842
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCAC 600

Qy       1843 ATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGT 1902
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGT 660
```

```
Qy       1903 GAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAA 1962
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAA 720

Qy       1963 TTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 770


RESULT 14
US-10-776-889-20
; Sequence 20, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Zhan-Bin
;  APPLICANT:  Kinney, Anthony
;  TITLE OF INVENTION: Seed Specific Promoters
;  FILE REFERENCE: BB1531 US NA
;  CURRENT APPLICATION NUMBER: US/10/776,889
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 20
;   LENGTH: 584
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-776-889-20

  Query Match             29.0%;  Score 584;  DB 19;  Length 584;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-64;
  Matches  584;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1429 ATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATAT 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATAT 60

Qy       1489 TTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTAC 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTAC 120

Qy       1549 AACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTA 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTA 180

Qy       1609 AGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGA 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGA 240

Qy       1669 AAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACA 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACA 300

Qy       1729 ACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTT 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTT 360

Qy       1789 TTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGT 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGT 420

Qy       1849 AATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGC 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGC 480

Qy       1909 GTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACA 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACA 540

Qy       1969 TAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 584
```

```
RESULT 15
US-10-776-889-21
; Sequence 21, Application US/10776889
; Publication No. US20040158052A1
; GENERAL INFORMATION:
;  APPLICANT: Liu, Zhan-Bin
;  APPLICANT:  Kinney, Anthony
;  TITLE OF INVENTION: Seed Specific Promoters
;  FILE REFERENCE: BB1531 US NA
;  CURRENT APPLICATION NUMBER: US/10/776,889
;  CURRENT FILING DATE:  2004-02-11
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Glycine max
US-10-776-889-21

  Query Match             21.1%;  Score 425;  DB 19;  Length 425;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-44;
  Matches  425;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1588 ATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTT 1647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTT 60

Qy       1648 CATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTT 1707
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTT 120

Qy       1708 ACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACT 1767
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACT 180

Qy       1768 TTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTA 1827
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTA 240

Qy       1828 TTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTA 1887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTA 300

Qy       1888 ACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGA 1947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGA 360

Qy       1948 CGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATA 2007
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATA 420

Qy       2008 CTTCA 2012
             |||||
Db        421 CTTCA 425


Search completed: September  2, 2005, 16:33:38
Job time : 989.534 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 22:19:32 ; Search time 6107.49 Seconds
(without alignments)
15962.684 Million cell updates/sec

Title: US-10-776-889-1
Perfect score: 2012
Sequence: 1 atcttaggcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	170.2	8.5	111861	9	AC069435	AC069435 Homo sapi
	2	170	8.4	349980	6	AX344555	AX344555 Sequence
c	3	168.2	8.4	90550	9	AL592166	AL592166 Human DNA
c	4	162	8.1	179668	2	CR626936	CR626936 Danio rer
c	5	161.2	8.0	14867	3	AE001398	AE001398 Plasmodiu
c	6	160.4	8.0	104992	2	AC005504	AC005504 Plasmodiu
c	7	160.4	8.0	169546	2	AC004157	AC004157 Plasmodiu
c	8	160.4	8.0	202169	2	AC149549	AC149549 Gorilla g
	9	160.4	8.0	250421	3	AE014849	AE014849 Plasmodiu
c	10	159.8	7.9	180096	2	CR626872	CR626872 Danio rer
c	11	159.6	7.9	178427	2	CR392006	CR392006 Danio rer
	12	158.6	7.9	120029	2	HSJ282H10	AL132672 Homo sapi
	13	158.4	7.9	143331	9	AC091214	AC091214 Homo sapi
c	14	158.4	7.9	226108	2	CR759967	CR759967 Danio rer
	15	158	7.9	129240	9	AC084128	AC084128 Homo sapi
	16	157.8	7.8	170627	2	AC125567	AC125567 Rattus no
c	17	157.8	7.8	181792	9	AC098822	AC098822 Homo sapi
c	18	157.6	7.8	125836	2	AC137629	AC137629 Homo sapi
c	19	156.8	7.8	79018	9	AL627107	AL627107 Human DNA
	20	156.8	7.8	135121	9	AC069525	AC069525 Homo sapi
	21	156	7.8	15548	6	AX347057	AX347057 Sequence
c	22	155.2	7.7	148937	2	CR361568	CR361568 Danio rer
	23	155.2	7.7	348174	3	CR382399	CR382399 Plasmodiu
c	24	154.6	7.7	172777	2	CR847809	CR847809 Danio rer
	25	154.6	7.7	176704	2	CR628364	CR628364 Danio rer
c	26	154.4	7.7	170143	2	AC048384	AC048384 Homo sapi
	27	154.2	7.7	171537	2	CR762390	CR762390 Danio rer
c	28	154.2	7.7	174198	2	CR762438	CR762438 Danio rer
	29	154	7.7	169800	2	CR847821	CR847821 Danio rer
	30	154	7.7	254050	3	PFA929358	AL929358 Plasmodiu
c	31	154	7.7	313050	3	PFA929352	AL929352 Plasmodiu
	32	153.6	7.6	186165	9	AC112187	AC112187 Homo sapi
	33	153.4	7.6	159618	9	AC096750	AC096750 Homo sapi
	34	153.2	7.6	157467	2	CR792423	CR792423 Danio rer
	35	153	7.6	149563	2	CR376750	CR376750 Danio rer
c	36	152.8	7.6	154044	2	CR786570	CR786570 Danio rer
	37	152.8	7.6	214709	2	CR749183	CR749183 Danio rer
c	38	152.8	7.6	228431	2	CR788237	CR788237 Danio rer
	39	152.6	7.6	210218	2	CR555304	CR555304 Danio rer
c	40	152.6	7.6	210871	2	CR759970	CR759970 Danio rer
	41	152.4	7.6	148075	9	AC110014	AC110014 Homo sapi
c	42	152.4	7.6	155106	9	AC104069	AC104069 Homo sapi
	43	151.8	7.5	251237	2	CR812792	CR812792 Danio rer
	44	151.6	7.5	178785	9	AC068139	AC068139 Homo sapi
c	45	151.6	7.5	227073	2	CR846086	CR846086 Danio rer

ALIGNMENTS

RESULT 1
AC069435/c
LOCUS AC069435 111861 bp DNA linear PRI 30-AUG-2002
DEFINITION Homo sapiens 3 BAC RP11-384L1 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC069435
VERSION AC069435.19 GI:20340410
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 111861)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,

Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 111861)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (30-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 111861)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (30-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 111861)

AUTHORS Worley,K.C.

TITLE Direct Submission

JOURNAL Submitted (30-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Apr 30, 2002 this sequence version replaced gi:20066162. INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.

FEATURES Location/Qualifiers

source 1. .111861
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-384L1"

misc_feature 1. .2005
/note="overlaps bases 165307. .167311 of clone AC073895"
/function="clone overlap"

repeat_region complement(1195. .1708)
/rpt_family="LTR33"

repeat_region 1882. .2418
/rpt_family="HAL1"

repeat_region 2563. .2586
/rpt_family="(TG)n"

repeat_region 2618. .2692
/rpt_family="(CATATA)n"

repeat_region 2935. .3049
/rpt_family="MIR"

repeat_region complement(3700. .3783)
/rpt_family="MIR"

repeat_region 4827. .4848
/rpt_family="AT_rich"

repeat_region 4981. .5016
/rpt_family="AT_rich"

repeat_region complement(5037. .5469)
/rpt_family="L1"

repeat_region complement(5721. .6053)
/rpt_family="L1"

repeat_region 6185. .6231
/rpt_family="AT_rich"

repeat_region 7100. .7297
/rpt_family="MIR"

repeat_region 8934. .9209
/rpt_family="MER102"

repeat_region 9667. .9968
/rpt_family="AluSx"

repeat_region complement(10412. .10660)
/rpt_family="LTR33"

repeat_region 10710. .10789
/rpt_family="MIR"

repeat_region 10941. .11011
/rpt_family="(TA)n"

repeat_region 11092. .11196
/rpt_family="T-rich"

repeat_region complement(11197. .11506)
/rpt_family="AluSx"

repeat_region 11507. .11599
/rpt_family="T-rich"

repeat_region 12304. .12379
/rpt_family="MER21B"

repeat_region 12405. .13097
/rpt_family="MER21B"

repeat_region 13104. .13257
/rpt_family="MLT1B"

repeat_region 13304. .13346
/rpt_family="(TA)n"

repeat_region 13849. .14101
/rpt_family="L1ME2"

repeat_region 14325. .14431
/rpt_family="LTR16B"

repeat_region 14446. .14536
/rpt_family="L1ME2"

repeat_region complement(15499. .15556)
/rpt_family="MIR"

repeat_region complement(15509. .15626)
/rpt_family="L2"

repeat_region complement(15700. .15761)
/rpt_family="MER92C"

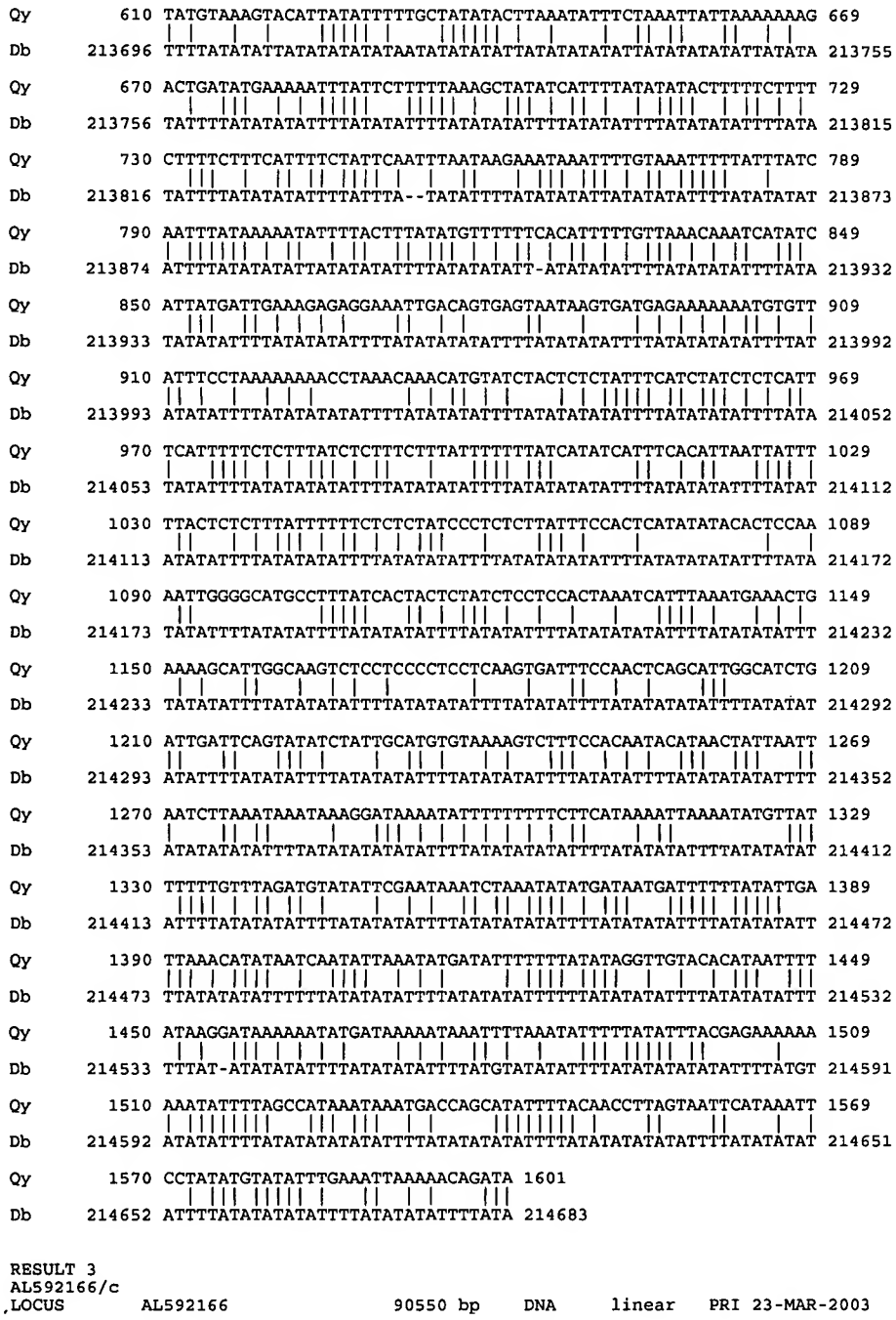
repeat_region 16137. .16200
/rpt_family="(TA)n"

repeat_region 17077. .17117
/rpt_family="AT_rich"

repeat_region 19849. .19949

```
                     /rpt_family="GA-rich"
     repeat_region   20599. .20717
                     /rpt_family="FLAM_C"
     repeat_region   20721. .20763
                     /rpt_family="(TAAA)n"
     repeat_region   21442. .21725
                     /rpt_family="MLT1A1"
     repeat_region   complement(21762. .22107)
                     /rpt_family="MLT1I"
     repeat_region   22360. .22395
                     /rpt_family="AT_rich"
     repeat_region   22559. .22817
                     /rpt_family="AluSx"
     repeat_region   complement(22927. .23015)

  Query Match             8.5%;  Score 170.2;  DB 9;  Length 111861;
  Best Local Similarity   48.6%;  Pred. No. 5.1e-11;
  Matches  551;  Conservative    0;  Mismatches  578;  Indels    4;  Gaps    3;

Qy        465 TACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTG 524
              || ||   |   |  ||| ||  | | |    ||||||||||     |||| | |||
Db      44345 TATATTTATATATATTTATATATACACACCATTTGTTTATATATAATATTATATATAAAT 44286

Qy        525 TCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTT 584
              | |   | |  |||||  ||||  |      ||  | | | | | ||    || ||
Db      44285 ATATATATTTTTATATATATATATAATATTATATATAAATATATATTTTTATATATTT 44226

Qy        585 CTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTA 644
              ||| |   | ||||   |      | | || | |    ||||  | |  ||||||  ||
Db      44225 ATATATAATATTTATATATATATTTTTTATATATTTATATATAAATATTATATATAAGTA 44166

Qy        645 AATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATA 704
              ||||| | |    |||     | |    | || | |    |   | ||| || |  |||
Db      44165 AATATATTTTTTATATATTTATATATTTTATTTATATTTATATATATTTATATATATATT 44106

Qy        705 TCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAA 764
              |   | ||| ||||  | ||  | | | || ||  || | |   |  |||||   |   |
Db      44105 TTTATATATTTATATATATTTATATATATTTTTATATATATTTATATATTTATATATTTA 44046

Qy        765 TAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTC 824
              || || || || || | | | |||  ||||||| | ||  | || | ||||| ||| |
Db      44045 TATATATTTTATATATATTTATATATATTTATATATATTATATATTATATATATTTATAT 43986

Qy        825 ACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTA 884
              | ||| |   | |   |    |||   | | ||  | |      | |||     |   ||
Db      43985 ATATTATATATTATATATATTTATATATTTTATATATATTATATATATTATATATATTTA 43926

Qy        885 ATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTA 944
              |      | | |   || | |  ||   || |  | |  ||     |   | ||| ||
Db      43925 TATATATTATATATATTTATATATATTTATATATATTATATATATTATATATTATATATA 43866

Qy        945 CTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTA 1004
              | | ||| | || ||| |   || |  || |     ||||| | |   |   || | ||
Db      43865 TTATATATATTATATATATTATATATATTTATATATATTATCTATATTATATATTATATA 43806

Qy       1005 TCATATCATTTCACATTAATTATTTTTACTC-TCTTTATTTTTTCTCTCTATCCCTCTCT 1063
              |  |   || |   ||  ||  | | || |  | | ||||| |    | |||   |
Db      43805 TATTTATATATATTATATATATTATATATTTATATATATTTATATATTATATATATTATA 43746

Qy       1064 TATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCC 1123
              ||||  |   || |||| | |  |  |          | ||| |    | ||| |
Db      43745 CATTTATATATATTTATATATTATATATATTTATATATTATATATTATATATTTATATAT 43686

Qy       1124 TCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGT 1183
              || || |  || || | |    |     ||     |          |    | |
Db      43685 ATTTATATATTAGATATATTATATATATTTATATATTATATATATTATATATTTATATAT 43626

Qy       1184 GATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGT 1243
                ||    | |    ||          ||| ||  | ||||| ||||  || | |  | |
Db      43625 ATTTATATATTATATATATTATATATATTTATATATTATATATATTATATATATTTATAT 43566

Qy       1244 CTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTT 1303
                ||      ||  |||  |||| ||  ||  || |||  || |   | ||||||  ||| |
```

```
Db      43565 ATTATATATATT-ATATATATTTATATATTATATATATTTATATATTTAAATATATTTAT 43507

Qy       1304 TTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAA 1363
                | ||    | |||| |      |||| | |  | || ||      |   ||| || |  |
Db      43506 ATATTATATATATTATATATATTTATATATATTATATATTATATATATTATATATATTTA 43447

Qy       1364 TATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTT 1423
              |||||  ||   ||| | ||||||   || |  ||||    ||||  | |||      | |
Db      43446 TATATATTATATATTATATATATTATATATATTTATATATTATATATATTATATATATAT 43387

Qy       1424 TTTATATAGGTTGTAC--ACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTT 1481
                ||||||   | ||   | |||  ||    |  ||| | | |   ||| |  || || |
Db      43386 AATATATAATATATATTTATATATATTATATATTATATATATTTATATATATTTATATAT 43327

Qy       1482 TAAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCAT 1541
              || |||| ||||||| ||     | |  | |||| |||    ||  ||| ||    |
Db      43326 TATATATATTTATATATATTTATATATTATATATATTATATATTTATATATATTTATAGA 43267

Qy       1542 ATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAA 1594
                |||      | | | ||  |||  |   ||||   |||| | || |  | |
Db      43266 TTTTATATGTATAAATTATATATATTTATATATAAAAATATATAAATTATATA 43214


RESULT 2
AX344555
LOCUS       AX344555               349980 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 6 from Patent WO0200932.
ACCESSION   AX344555
VERSION     AX344555.1  GI:18492441
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of known genetic parameters within the mhc
  JOURNAL   Patent: WO 0200932-A 6 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .349980
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo
                     sapiens)~Original length of seq 1: 3.673778 <223>~split as
                     follows.~seq 01 0.000.001 TO 0.349.980~seq 02 0.300.001
                     649.980~seq 03 600.001 949.980~seq 04 900.001
                     1.249.980~seq 05 1.200.001 1.549.980~seq 06 1.500.001
                     1.849.980~seq 07 1.800.001 2.149.980~seq 08 2.100.001
                     2.449.980~seq 09 2.400.001 2.749.980~seq 10 2.700.001
                     3.049.980~seq 11 3.000.001 3.349.980~seq 12 3.300.001
                     3.649.980~seq 13 3.600.001 3.673.778 <223>~Original length
                     of seq 2: 3.673778 <223>~split as follows.~seq 14
                     0.000.001 TO 0.349.980~seq 15 0.300.001 649.980~seq 16
                     600.001 949.980~seq 17 900.001 1.249.980~seq 18 1.200.001
                     1.549.980~seq 19 1.500.001 1.849.980~seq 20 1.800.001
                     2.149.980~seq 21 2.100.001 2.449.980~seq 22 2.400.001
                     2.749.980~seq 23 2.700.001 3.049.980~seq 24 3.000.001
                     3.349.980~seq 25 3.300.001 3.649.980~seq 26 3.600.001
                     3.673.778"
ORIGIN

  Query Match             8.4%;  Score 170;  DB 6;  Length 349980;
  Best Local Similarity   48.8%;  Pred. No. 3.6e-11;
  Matches  543;  Conservative    0;  Mismatches  565;  Indels    4;  Gaps    3;

Qy        490 AGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTA 549
              | | || ||·| |||||| | ||     ||||      |    | ||| ||||| ||
Db     213576 ATATATTTTATATATATTTTATATATTTTATATATATTTTATATATATTTTATATATTTT 213635

Qy        550 ATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCA 609
              ||   ||| |||  | | | |    | ||||   | ||||||  |  |||| | |||  |
Db     213636 ATATATATTTTATATATATTTTATATATATTTTATATATTTTATATATATTTTATATATA 213695
```

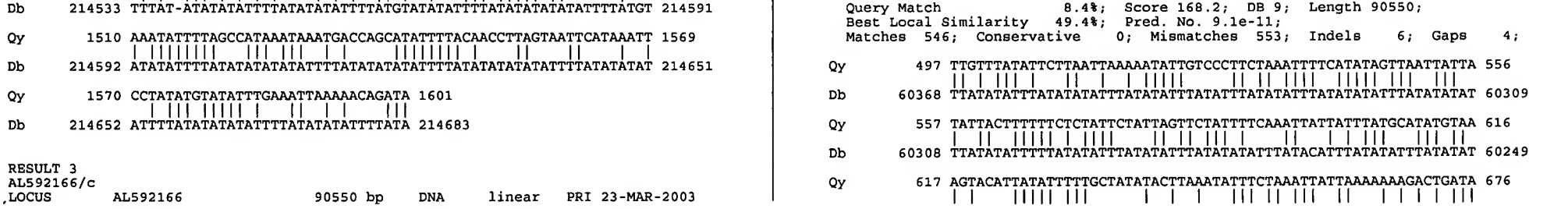


RESULT 3
AL592166/c
LOCUS AL592166 90550 bp DNA linear PRI 23-MAR-2003

DEFINITION Human DNA sequence from clone RP11-269F19 on chromosome 1, complete sequence.
ACCESSION AL592166
VERSION AL592166.16 GI:29165241
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 90550)
AUTHORS Wallis,J.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2003) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT On Mar 23, 2003 this sequence version replaced gi:18476714.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr1
RP11-269F19 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
-------------- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
FEATURES Location/Qualifiers
source 1. .90550
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-269F19"
/clone_lib="RPCI-11.1"
ORIGIN

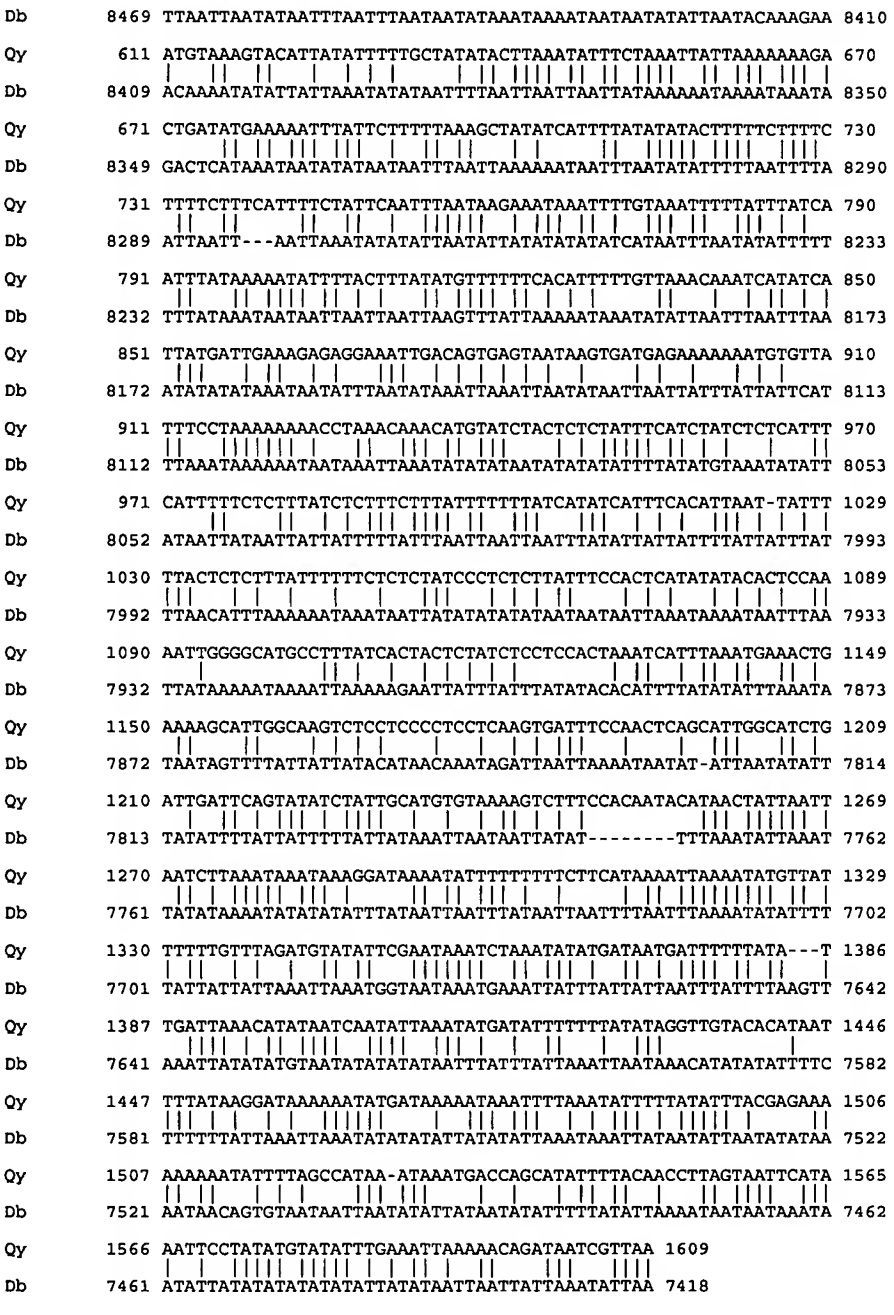
Query Match 8.4%; Score 168.2; DB 9; Length 90550;
Best Local Similarity 49.4%; Pred. No. 9.1e-11;
Matches 546; Conservative 0; Mismatches 553; Indels 6; Gaps 4;



```
Db      60248 ATTTATATATATATTTATATATTTATATATATTTTTTATATTTATATTTATATATTTATA 60189

Qy        677 TGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCT 736

Db      60188 TATATATTTATATTTATATATTGTTATATATTTTTATATATATTTATATTTATATATTGT 60129

Qy        737 TTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTAT 796

Db      60128 TATATATTTTTATATATTTATATATTGTTATATATT---TATATATTTATATATTGTTAT 60072

Qy        797 AAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGA 856

Db      60071 ATATTTTTATATATATTTATACATTTATATATTTATATATATTTATATATTTATATATAT 60012

Qy        857 TTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCT 916

Db      60011 TTTTTATATTTATATTTATATATTTATATATATTTATATTTATATATTGTTATATATTTT 59952

Qy        917 AAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTT 976

Db      59951 TATATATATTTATATTTATATATTGTTATATATTTTTATATATATTTATATATTGTTATA 59892

Qy        977 TCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCT 1036

Db      59891 TATTTTTATATATTTATATATTTATATATTTATATATATTTATATATATATTTATATATA 59832

Qy       1037 CTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGG 1096

Db      59831 TTTATATATTTATATATATTTATAT-ATATTTATATATATTTATATATATTTATATATTT 59773

Qy       1097 GCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCA 1156

Db      59772 ATATATATTTATATATATTTATATATTTATATTTATATATTTATATATTTATATATATTT 59713

Qy       1157 TTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATT 1216

Db      59712 ATATATATTTATATATTTATATATTTATATATATTTATATATATTTATATATTTATATAT 59653

Qy       1217 CAGTATATCTA-TTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTT 1275

Db      59652 ATTTATATATATTTATATATTTATATATATTTATATATTTATATATATTTTTATATATTT 59593

Qy       1276 AAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTG 1335

Db      59592 TTATATATTTATATTTATATATATTTATATATTTATATATATTTTTTATATATTTATATA 59533

Qy       1336 TTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAAC 1395

Db      59532 TTCATATATATTCATAAATACTTATATTTATTTATTTATATATGTATATATGTATATATT 59473

Qy       1396 ATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGG 1455

Db      59472 TATATATTTATATATATTTATATATTTATATATATTTTTTATATATTTATATATTCATAT 59413

Qy       1456 ATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAATAT 1515

Db      59412 ATATTCATAAATACTTATATTTA-TTTATATATTTATATATTTATATATATTTATATATT 59354

Qy       1516 TTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATA 1575

Db      59353 TATATATTTATATATATTTTTATATATTTATATATTTATATATATTTTTATATATTTATA 59294

Qy       1576 TGTATATTTGAAATTAAAAACAGAT 1600

Db      59293 TATATTTATATATATTTATATATAT 59269


RESULT 4
CR626936/c
LOCUS       CR626936              179668 bp    DNA     linear   HTG 09-AUG-2004
DEFINITION  Danio rerio clone DKEY-107I22, WORKING DRAFT SEQUENCE, 5 unordered
            pieces.
ACCESSION   CR626936
VERSION     CR626936.3  GI:51090206
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 179668)
  AUTHORS   Burton,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-AUG-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 9, 2004 this sequence version replaced gi:50896321.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK107I22
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 177410 bases at least Q40
            Consensus quality: 177759 bases at least Q30
            Consensus quality: 178052 bases at least Q20
            Insert size: 179268; sum-of-contigs
            Insert size: 173635; 3.1% error; agarose-fp
            Quality coverage: 9.22x in Q20 bases; sum-of-contigs Quality
            coverage: 9.90x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     5950: contig of 5950 bp in length
            *     5951     6050: gap of 100 bp
            *     6051    72517: contig of 66467 bp in length
            *    72518    72617: gap of 100 bp
            *    72618    87634: contig of 15017 bp in length
            *    87635    87734: gap of 100 bp
            *    87735   130094: contig of 42360 bp in length
            *   130095   130194: gap of 100 bp
            *   130195   179668: contig of 49474 bp in length.
FEATURES             Location/Qualifiers
     source          1. .179668
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-107I22"
                     /clone_lib="DanioKey"
     misc_feature    1. .5950
                     /note="assembly_fragment:00021
                     fragment_chain:1"
     misc_feature    6051. .72517
                     /note="assembly_fragment:01524
                     fragment_chain:1"
     misc_feature    72618. .87634
                     /note="assembly_fragment:00072
                     fragment_chain:1"
     misc_feature    87735. .130094
                     /note="assembly_fragment:00264"
     misc_feature    130195. .179668
                     /note="assembly_fragment:00871.0"
ORIGIN

  Query Match             8.1%;  Score 162;  DB 2;  Length 179668;
  Best Local Similarity   45.0%;  Pred. No. 3.6e-10;
```

```
  Matches  498;  Conservative    0;  Mismatches  605;  Indels    3;  Gaps    3;

Qy        498 TGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTAT 557
              | | |||||   ||  || | ||||      | || || |  |||||  ||  |||
Db      73226 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 73167

Qy        558 ATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAA 617
              ||  | | | | |||     ||  | ||| |  | ||   ||| |||  |||| || |
Db      73166 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 73107

Qy        618 GTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATAT 677
              |||||  | |  ||||||  || |||| | || ||   || | | | |   ||||
Db      73106 TATATATATATATAT-ATATATATATATATATATATATATATATATATATATATATAT 73048

Qy        678 GAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTT 737
              | |  |   | | | || |  |||||   | ||||||||  | |   | | | |   |
Db      73047 ATATATATATATATATATATATATATATATATATATATATATAGATATATATATAT 72988

Qy        738 TCATTTTCTATTCAATTTAATAAGAAATAAAT-TTTGTAAATTTTTATTTATCAATTTAT 796
              || |     |  || ||   | | ||| || | | || || | ||| |||  || |||
Db      72987 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 72928

Qy        797 AAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGA 856
              | | |||| |   | |||||| | | |  | || | | |   |   |  ||||  |||
Db      72927 ATACATATATATATATATATATATATATATATATATATATATATATATATATATATAT 72868

Qy        857 TTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCT 916
              | |      | ||  | | |   |||  |   | | | | || | | || |
Db      72867 ATATATATATATATATATATATATATATATATATATATATATATATATATATCATATATA 72808

Qy        917 AAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTT 976
              | | | |  || | | | ||  || ||     |   | || ||| | |   |  || |
Db      72807 TATATATATATATATATATAGATATATATACATATATATATATATATATATATATATA 72748

Qy        977 TCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCT 1036
              | | | ||| | | | | ||| | | |||       || |      ||  ||| | ||
Db      72747 TATATATATATATATATATATATATATATATATATATATATATATACATATATATATATA 72688

Qy       1037 CTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGG 1096
                |   | | |   | |||   | || ||| |  |   |||||  | |  |    ||
Db      72687 TGTATATGTATAAATATATATATATCATATATATATATATATAGATATACATCTATATAT 72628

Qy       1097 GCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCA 1156
              ||   |
Db      72627 ATATATATACNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 72568

Qy       1157 TTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATT 1216
                                                                |   |   |
Db      72567 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNATATTTTAT 72508

Qy       1217 CAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTA 1276
                ||||| |||   || | || |  | |||  | |       |  ||  ||   ||
Db      72507 ATATATATATATATAATTTATATATATATTTATATATATATATTTTTATATATATATATA 72448

Qy       1277 AATAAATAAAGGATAAAATATTTTTTTTTCTTCATA-AAATTAAAATATGTTATTTTTTG 1335
              ||| ||| |     | |||| | | | | |  ||| | ||  | |||| |   ||| |
Db      72447 TATATATATATATATATATATATATATATATATATATATATATATATATATATATTTATA 72388

Qy       1336 TTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAAC 1395
              | || || |||||    ||| || || ||||||    ||   | | |||||  ||  |
Db      72387 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 72328

Qy       1396 ATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGG 1455
                ||  ||||| |    |||| ||| ||||||  |          || | |   |
Db      72327 TATATATATATATTTATATATATATATTTATATATATATATATATATATATATATATA 72268

Qy       1456 ATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAATAT 1515
                | | ||||    | |   | | || |||| | ||||| ||     | | | |
Db      72267 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 72208

Qy       1516 TTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATA 1575
              | ||   ||| ||| ||    |      | || |    ||   ||  ||| ||   ||||
Db      72207 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 72148

Qy       1576 TGTATATTTGAAATTAAAAACAGATA 1601
                | ||||| |  |  || | | | |||
Db      72147 TATATATATATATATATATATATATA 72122


RESULT 5
AE001398/c
LOCUS       AE001398               14867 bp    DNA     linear   INV 04-OCT-2002
DEFINITION  Plasmodium falciparum 3D7 chromosome 2 section 35 of 73 of the
            complete sequence.
ACCESSION   AE001398 AE001362
VERSION     AE001398.1  GI:3845197
KEYWORDS    .
SOURCE      Plasmodium falciparum 3D7
  ORGANISM  Plasmodium falciparum 3D7
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 14867)
  AUTHORS   Gardner,M.J., Tettelin,H., Carucci,D.J., Cummings,L.M., Aravind,L.,
            Koonin,E.V., Shallom,S., Mason,T., Yu,K., Fujii,C., Pederson,J.,
            Shen,K., Jing,J., Aston,C., Lai,Z., Schwartz,D.C., Pertea,M.,
            Salzberg,S., Zhou,L., Sutton,G.G., Clayton,R., White,O.,
            Smith,H.O., Fraser,C.M., Adams,M.D., Venter,J.C. and Hoffman,S.L.
  TITLE     Chromosome 2 sequence of the human malaria parasite Plasmodium
            falciparum
  JOURNAL   Science 282 (5391), 1126-1132 (1998)
  MEDLINE   99021743
   PUBMED   9804551
REFERENCE   2  (bases 1 to 14867)
  AUTHORS   Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
            Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
            James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
            Kyes,S., Chan,M.-S., Nene,V., Shallom,S.J., Suh,B., Peterson,J.,
            Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
            Mather,M.W., Vaidya,A.B., Martin,D.M.A., Fairlamb,A.H.,
            Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
            Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
            Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
            Barrell,B.
  TITLE     Genome sequence of the human malaria parasite Plasmodium falciparum
  JOURNAL   Nature 419, 498-511 (2002)
REFERENCE   3  (bases 1 to 14867)
  AUTHORS   Gardner,M.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-OCT-2002) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1. .14867
                     /organism="Plasmodium falciparum 3D7"
                     /mol_type="genomic DNA"
                     /isolate="3D7"
                     /db_xref="taxon:36329"
                     /chromosome="2"
     repeat_region   complement(<1. .20)
                     /note="AT_rich"
                     /rpt_type=tandem
     repeat_region   complement(65. .137)
                     /note="AT_rich"
                     /rpt_type=tandem
     repeat_region   complement(178. .211)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
     repeat_region   307. .327
                     /rpt_type=tandem
                     /rpt_unit="(a)n"
     repeat_region   complement(328. .348)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
     repeat_region   complement(378. .423)
                     /note="AT_rich"
                     /rpt_type=tandem
     repeat_region   complement(479. .503)
                     /note="AT_rich"
```

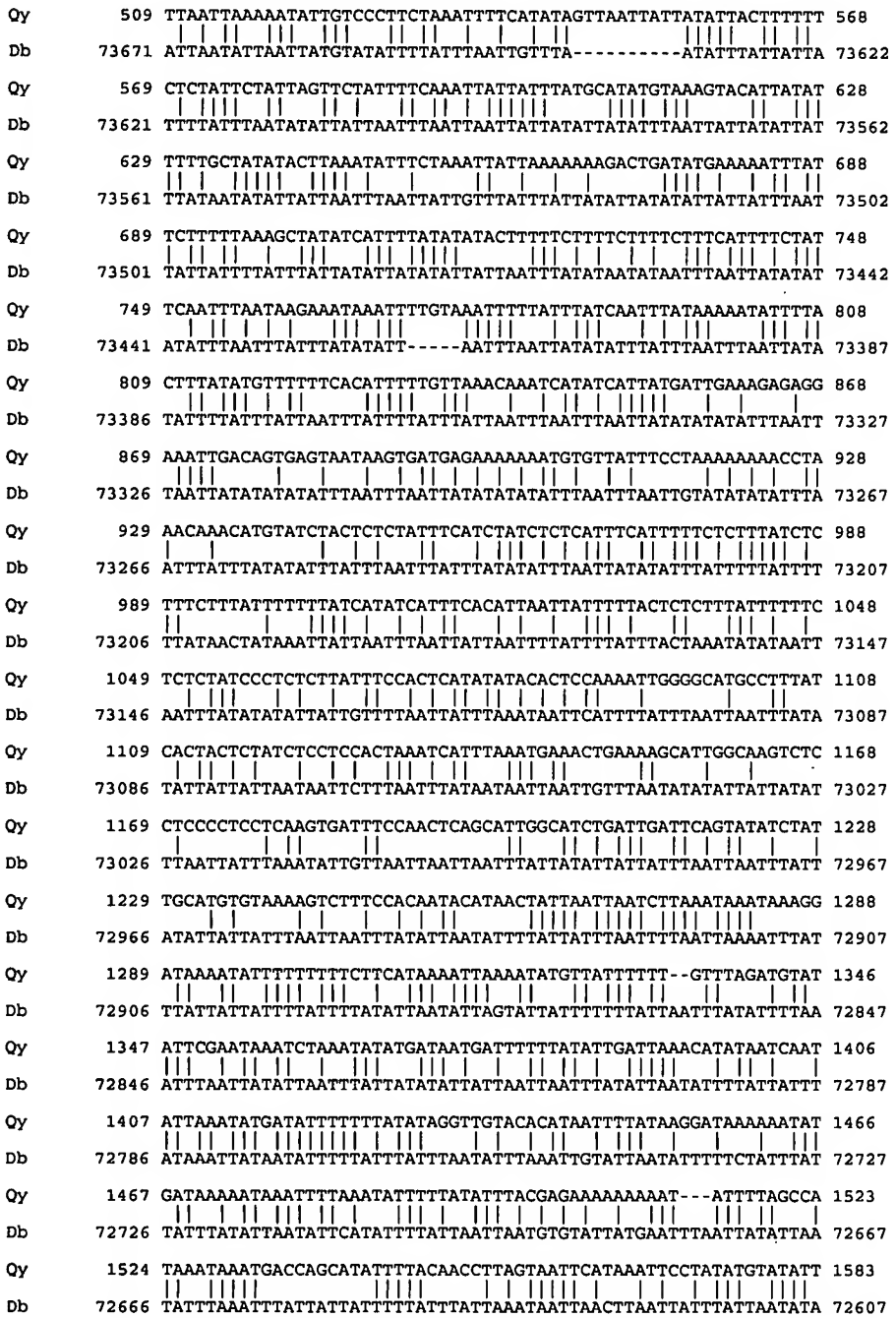
```
                     /rpt_type=tandem
repeat_region        complement(513. .539)
                     /rpt_type=tandem
                     /rpt_unit="(a)n"
repeat_region        577. .633
                     /rpt_type=tandem
                     /rpt_unit="(a)n"
repeat_region        complement(714. .747)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        841. .867
                     /rpt_type=tandem
                     /rpt_unit="(a)n"
repeat_region        complement(888. .918)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(1006. .1049)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(1075. .1120)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(1140. .1283)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        1294. .1346
                     /rpt_type=tandem
repeat_region        1298. .1346
                     /rpt_type=tandem
                     /rpt_unit="(tataa)n"
repeat_region        complement(1423. .1531)
                     /note="AT_rich"
                     /rpt_type=tandem
gene                 complement(<1570. .>2424)
                     /locus_tag="PFB0490c"
mRNA                 complement(<1570. .>2424)
                     /locus_tag="PFB0490c"
CDS                  complement(1570. .2424)
                     /locus_tag="PFB0490c"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAC71887.1"
                     /db_xref="GI:3845198"
                     /translation="MKEKNEKIMDYLSCPLDDVVDREKKSGKNSLLKSSSTKKSDYKK
                     SSIFSKKRDSHKKGSSFRGRRSGFINRKSGSFKKPYYNNRLINKNYNNYKGRNFHNGR
                     DNFKGRTGSFGSRVFDNRKGSFKKRFISNRNKSSVKSYRGNGSNKMGRKSFNKAPTSR
                     TVVTKRLNNYKTVSAPVKKFNNLNISLYRKNRTFALNTKRSKPVGTIKSSVPRKRIKK
                     GLKKGSLKSKTRKSTSGSKFKPLNKYFLSKIKIVTSLNKIPSPLKEQKNTEVNLPESL
                     NNATTKKN"
repeat_region        complement(2150. .2188)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(2446. .2468)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
repeat_region        2504. .2656
                     /rpt_type=tandem
                     /rpt_unit="(tataa)n"
repeat_region        2684. .2739
                     /rpt_type=tandem
repeat_region        2693. .2736
                     /rpt_type=tandem
repeat_region        complement(2697. .2732)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
repeat_region        complement(2757. .2845)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(2856. .2880)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
repeat_region        2884. .2953
                     /rpt_type=tandem
                     /rpt_unit="(tataa)n"
repeat_region        2888. .2924
                     /rpt_type=tandem
repeat_region        complement(2977. .3167)
                     /rpt_type=tandem
                     /rpt_unit="(taaa)n"
repeat_region        3047. .3132
                     /rpt_type=tandem
repeat_region        complement(3047. .3132)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
repeat_region        complement(3258. .3420)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(3450. .3514)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
repeat_region        complement(3532. .3559)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(3689. .3720)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(4024. .4114)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        4107. .4146
                     /rpt_type=tandem
repeat_region        complement(4115. .4146)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
repeat_region        4227. .4261
                     /rpt_type=tandem
repeat_region        complement(4227. .4257)
                     /rpt_type=tandem
                     /rpt_unit="(ta)n"
repeat_region        complement(4348. .4386)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(4403. .4450)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(4545. .4785)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(4937. .5001)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(5102. .5191)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(5246. .5275)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(5437. .5484)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(5554. .5577)
                     /note="AT_rich"
                     /rpt_type=tandem
repeat_region        complement(6302. .6350)
                     /note="AT_rich"

  Query Match             8.0%;  Score 161.2;  DB 3;  Length 14867;
  Best Local Similarity   50.2%;  Pred. No. 1e-09;
  Matches  564;  Conservative    0;  Mismatches  543;  Indels   17;  Gaps    6;

Qy       491 GAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAA 550
             |||||| |  ||||  | | |||  | | ||   |   |   | | | | |||   || |
Db      8529 GAAATAATTATTATTATATAAATATATATTAAAATAAATGAAATAATATTATAATTTTTA 8470

Qy       551 TTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCAT 610
             ||| |  ||    |||   |  ||  ||| | |   ||  | | ||  | ||  |   |
```

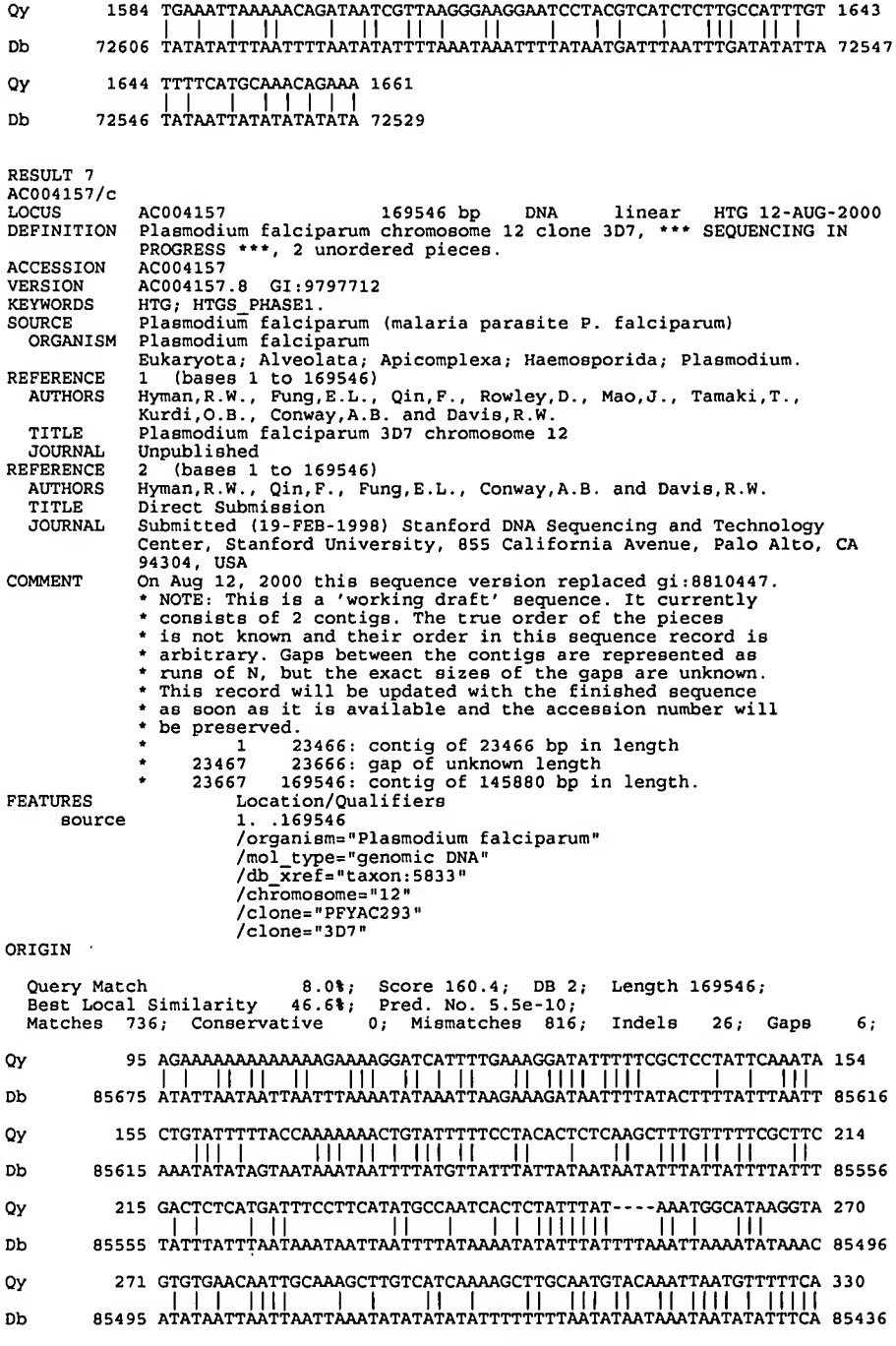


RESULT 6
AC005504/c

LOCUS AC005504 104992 bp DNA linear HTG 01-APR-1999
DEFINITION Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS ***, 3 unordered pieces.
ACCESSION AC005504
VERSION AC005504.3 GI:4558584
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 104992)
AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B., Conway,A.B. and Davis,R.W.
TITLE Plasmodium falciparum 3D7 chromosome 12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 104992)
AUTHORS Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT On Apr 2, 1999 this sequence version replaced gi:4337172.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 58642: contig of 58642 bp in length
* 58643 58842: gap of unknown length
* 58843 91011: contig of 32169 bp in length
* 91012 91211: gap of unknown length
* 91212 104992: contig of 13781 bp in length.
FEATURES Location/Qualifiers
source 1. .104992
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
ORIGIN

Query Match 8.0%; Score 160.4; DB 2; Length 104992;
Best Local Similarity 46.6%; Pred. No. 6.5e-10;
Matches 736; Conservative 0; Mismatches 816; Indels 26; Gaps 6;





RESULT 7
AC004157/c
LOCUS AC004157 169546 bp DNA linear HTG 12-AUG-2000
DEFINITION Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN PROGRESS ***, 2 unordered pieces.
ACCESSION AC004157
VERSION AC004157.8 GI:9797712
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE 1 (bases 1 to 169546)
AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T., Kurdi,O.B., Conway,A.B. and Davis,R.W.
TITLE Plasmodium falciparum 3D7 chromosome 12
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 169546)
AUTHORS Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT On Aug 12, 2000 this sequence version replaced gi:8810447.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 23466: contig of 23466 bp in length
* 23467 23666: gap of unknown length
* 23667 169546: contig of 145880 bp in length.
FEATURES Location/Qualifiers
source 1. .169546
/organism="Plasmodium falciparum"
/mol_type="genomic DNA"
/db_xref="taxon:5833"
/chromosome="12"
/clone="PFYAC293"
/clone="3D7"
ORIGIN

Query Match 8.0%; Score 160.4; DB 2; Length 169546;
Best Local Similarity 46.6%; Pred. No. 5.5e-10;
Matches 736; Conservative 0; Mismatches 816; Indels 26; Gaps 6;

```
Qy       331 TGCCTT--TCAAAATTATCTGCACCCCCTAGCTATTAATCTAACATCTAAGTAAGGCTAG 388
Db     85435 CATTTTAATTAAAATAAAAATAACCATTTATTAATTAACTTAATTAATATATAAAATAAA 85376

Qy       389 TGAATTTTTTCGAATAGTCATGCAGTGCATTAATTTCCCCGTGACTATTTTGGCTTTGAC 448
Db     85375 ATAATTTAATTGTGTAATTAAATTAAATATAAAACATTTATTAATAATTAATTATATATA 85316

Qy       449 TCCAACACTGGCCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTC 508
Db     85315 ATATTATATATTATCTTAAATTAATTAATTTTTTTAATTATTTTAATATAATAATTATTT 85256

Qy       509 TTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTT 568
Db     85255 ATTAATATTAATTATGTATATTTTATTTAATTGTTTA----------ATATTTATTATTA 85206

Qy       569 CTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATAT 628
Db     85205 TTTTATTTAATATATTATTAATTTAATTAATTATTATATTATATTTAATTATTATATTAT 85146

Qy       629 TTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTAT 688
Db     85145 TTATAATATATTATTAATTTAATTATTGTTTATTTATTATATTATATATTATTATTTAAT 85086

Qy       689 TCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTAT 748
Db     85085 TATTATTTTATTTATTATATTATATATTATTAATTTATATAATATAATTTAATTATATAT 85026

Qy       749 TCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTA 808
Db     85025 ATATTTAATTTATTTATATATT-----AATTTAATTATATATTTATTTAATTTAATTATA 84971

Qy       809 CTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGG 868
Db     84970 TATTTTATTTATTAATTTATTTTATTTATTAATTTAATTTAATTATATATATATTTAATT 84911

Qy       869 AAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTA 928
Db     84910 TAATTATATATATATTTAATTTAATTATATATATATTTAATTTAATTGTATATATATTTA 84851

Qy       929 AACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTC 988
Db     84850 ATTTATTTATATATTTATTTAATTTATTTATATATTTAATTATATATTTATTTTTATTTT 84791

Qy       989 TTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTC 1048
Db     84790 TTATAACTATAAATTATTAATTTAATTATTAATTTTTATTTTATTTACTAAATATATAATT 84731

Qy      1049 TCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTAT 1108
Db     84730 AATTTATATATATTATTGTTTTAATTATTTAAATAATTCATTTTATTTAATTAATTTATA 84671

Qy      1109 CACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTC 1168
Db     84670 TATTATTATTAATAATTCTTTAATTTATAATAATTAATTGTTTAATATATATTATTATAT 84611

Qy      1169 CTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTAT 1228
Db     84610 TTAATTATTTAAATATTGTTAATTAATTAATTTATTATATTATTATTTAATTAATTTATT 84551

Qy      1229 TGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGG 1288
Db     84550 ATATTATTATTTAATTAATTTATATTAATATTTTATTATTTAATTTTAATTAAAATTTAT 84491

Qy      1289 ATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTT--GTTTAGATGTAT 1346
Db     84490 TTATTATTATTTTATTTTATATTAATATTAGTATTATTTTTTTATTAATTTATATTTTAA 84431

Qy      1347 ATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAAT 1406
Db     84430 ATTTAATTATATTAATTTATTATATATTATTAATTAATTTATATTAATATTTTATTATTT 84371

Qy      1407 ATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATAT 1466
Db     84370 ATAAATTATAATATTTTTATTTATTTAATATTTAAATTGTATTAATATTTTTCTATTTAT 84311

Qy      1467 GATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAAT---ATTTTAGCCA 1523
Db     84310 TATTTATATTAATATTCATATTTTATTAATTAATGTGTATTATGAATTTAATTATATTAA 84251

Qy      1524 TAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATT 1583
Db     84250 TATTTAAATTTATTATTATTTTTATTTATTAAATAATTAACTTAATTATTTATTAATATA 84191

Qy      1584 TGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGT 1643
Db     84190 TATATATTTAATTTTAATATATTTTAAATAAATTTTATAATGATTTAATTTGATATATTA 84131

Qy      1644 TTTTCATGCAAACAGAAA 1661
Db     84130 TATAATTATATATATATA 84113
```

RESULT 8
AC149549/c

```
LOCUS       AC149549              202169 bp    DNA     linear   HTG 30-JUL-2004
DEFINITION  Gorilla gorilla gorilla clone CH255-73E2, WORKING DRAFT SEQUENCE, 6
            ordered pieces.
ACCESSION   AC149549
VERSION     AC149549.2  GI:50838874
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      Gorilla gorilla gorilla (lowland gorilla)
  ORGANISM  Gorilla gorilla gorilla
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE   1  (bases 1 to 202169)
  AUTHORS   Walker,M.L., Osoegawa,K., Friedman,C., de Jong,P.J., Trask,B.J.,
            Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
            Coleman,H., Daki,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
            Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
            Karlins,E., Kim,H., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-Q.,
            Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
            Masiello,C., Maskeri,B., McDowell,J., Mullikin,J.C., Paguirigan,C.,
            Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddix-Dugue,N.,
            Schandler,K., Schueler,M.G., Shah,K., Sison,C., Stantripop,S.,
            Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D.,
            Young,A. and Green,E.D.
  TITLE     NISC Comparative Sequencing Initiative
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 202169)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE   3  (bases 1 to 202169)
  AUTHORS   Green,E.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2004) NIH Intramural Sequencing Center, 8717
            Grovemont Circle, Gaithersburg, MD 20877, USA
COMMENT     On Jul 30, 2004 this sequence version replaced gi:48427674.
            -------------- Genome Center
                 Center: NIH Intramural Sequencing Center
                 Center code: NISC
                 Web site: http://www.nisc.nih.gov
                 Contact: nisc_zoo@nhgri.nih.gov
            -------------- Project Information
                 Center project name: eqn
                 Center clone name: 073E02

            The sequence data in this record represents an 'enhanced'
            version of a Phase 2 submission. Specifically, the indicated
            order and orientation of each sequence contig has been
            established using one or more of the following: read-pair
            data from individual subclones, overlaps with neighboring
            clones, alignment with available reference sequence (e.g.,
```

```
          human), and/or confirmation by PCR testing. In addition,
          the sequence assembly is based on at least 8X average
          coverage in Q20 bases and has been reviewed to rule out
          gross misassemblies, the low-quality ends of sequence
          contigs have been trimmed away, and each base is associated
          with a Phrap-derived quality score.
          ---------- Summary Statistics
               Sequencing vector: plasmid; n/a; 100% of reads
               Chemistry: Dye-terminator Big Dye; 100% of reads
               Assembly program: Phrap; version 0.990319
               Consensus quality: 200572 bases at least Q40
               Consensus quality: 201163 bases at least Q30
               Consensus quality: 201502 bases at least Q20
               Insert size: 176000; agarose-fp
               Insert size: 201669; sum-of-contigs
               Quality coverage: 10.25x in Q20 bases; agarose-fp
               Quality coverage: 8.94x in Q20 bases; sum-of-contigs
          ----------.
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 6 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * provided by the submittor.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          *        1    15312: contig of 15312 bp in length
          *    15313    15412: gap of unknown length
          *    15413    73350: contig of 57938 bp in length
          *    73351    73450: gap of unknown length
          *    73451    80732: contig of 7282 bp in length
          *    80733    80832: gap of unknown length
          *    80833   133153: contig of 52321 bp in length
          *   133154   133253: gap of unknown length
          *   133254   159871: contig of 26618 bp in length
          *   159872   159971: gap of unknown length
          *   159972   202169: contig of 42198 bp in length.
FEATURES             Location/Qualifiers
     source          1. .202169
                     /organism="Gorilla gorilla gorilla"
                     /mol_type="genomic DNA"
                     /sub_species="gorilla"
                     /db_xref="taxon:9595"
                     /clone="CH255-73E2"
                     /clone_lib="CH255"
                     /note="BAC resource: http://bacpac.chori.org/"
     misc_feature    1. .15312
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    15413. .73350
                     /note="assembly_fragment"
     misc_feature    73451. .80732
                     /note="assembly_fragment"
     misc_feature    80833. .133153
                     /note="assembly_fragment"
     misc_feature    133254. .159871
                     /note="assembly_fragment"
     misc_feature    159972. .202169
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             8.0%;  Score 160.4;  DB 2;  Length 202169;
  Best Local Similarity   49.6%;  Pred. No. 5.2e-10;
  Matches  551;  Conservative    0;  Mismatches  546;  Indels   13;  Gaps    5;

Qy        496 ATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATT 555
              | | |||||||   |   |  | | ||      || || ||||  || ||  || ||||
Db      11426 AATATTTATATATATTTCTGTATAAATATATATTTATATATTTCTATTTAAATATTTATA 11367

Qy        556 ATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTA 615
              |   | | |   ||||    ||  | ||| || ||||  |||| |||| |||| ||
Db      11366 TATATTTATATATAAATATTTATATATATTTATATTTAAATATTTATATATGAATATATA 11307

Qy        616 AAGTACATTATATTT---TTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACT 672
              |  |  ||||| |   |   ||||||   | |||| | |||||   || | ||| | |
Db      11306 TATAAATATATATATAAATATTTATATATGAATATATATATAAATATATATATAAATATT 11247

Qy        673 GATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTT 732
              ||||   ||  |   | |   || | ||||    |||||||| |  | |     | |
Db      11246 TATATATGAATATATATATAAATATATATATAAATATTTATATAAATATATATAAATATA 11187

Qy        733 TTCTTTCAT-TTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAA 791
              |  |  || ||| |||   | |  |||   ||  |  | | ||||| ||||| ||| ||
Db      11186 TATATAAATATTTATATATGAATATATATATAAATATATATATAAATATTTATATATAAA 11127

Qy        792 TTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCAT 851
              | |||| ||||||||   | || || | | | | | | | ||   |   || | |
Db      11126 TATATATAAATATTTATATATAAATATATATAAATAGATATATAAATATATATATCTATT 11067

Qy        852 TATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTAT 911
              |||   | ||        | || | |  | | ||| | |  | | ||| | |   ||
Db      11066 TATATATAAATATATAAATATATATATCTATTTATATATAAATATATAAATAAATTAAAT 11007

Qy        912 TTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTC 971
              |  |||| |||     |  ||   | ||| ||      |  |  ||| |    |
Db      11006 ATATTAAATAAATAAATATAAATATTATATATATAAATACATATAAATATATAAATATAT 10947

Qy        972 ATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTT 1031
              ||   | | | | | | | | | ||| | ||    |||     |  | ||  ||  || |
Db      10946 ATAAATATATATTTATATATTTGTATATATTATATATAAATATATAAATACATATTTATA 10887

Qy       1032 ACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAA 1091
              | | ||||| | ||     | |      |  ||| |  |   || ||||||     | |
Db      10886 TATATATTTATATATTTATATTTTTATATTTATATATTTATATATTTATACATATTTATA 10827

Qy       1092 TTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAA 1151
              |      ||   | |||   ||  ||| |     | || ||  ||  |   |  | |  |
Db      10826 TATTTATATTTATATATTTGTATATTATATATATAAATATATATTTATATATTATATATA 10767

Qy       1152 AAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGAT 1211
              |    |         | |    |  |  |    | |  |          ||    || | |
Db      10766 TAATTATA-------TATTTATATATTATATATTTATATATTAATATATAAATATATAAA 10714

Qy       1212 TGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAA 1271
              |    | | ||| |    |         |  | |     |     ||||  ||||||  ||| |
Db      10713 TATATAAATATGTAACTATATATTTATATATAAATATATAAATATGTAACTATGTATTTA 10654

Qy       1272 TCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTT 1331
              | |  ||  |   ||  ||  ||  | | |||   |  || ||  |||||||| | |
Db      10653 TATATAAATATATAAATATGTAACTATATATTTATATATAAATATATAAATATGTAACTA 10594

Qy       1332 TTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATT 1391
              | | |||| || || ||     | | || ||| |||||  | ||    |  |||||  ||
Db      10593 TATATTTATATATAAATATATAAATATGTAACTATATATTTATATATAAATATATAAATA 10534

Qy       1392 AAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTAT 1451
              |     ||   |||  |  |  ||||   | |||| |    | || | |||| | |||
Db      10533 TGTAACTATATATTTATGTATAAATATATAAATATATA-AATATATAAATATAAATATAT 10475

Qy       1452 AAGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAA 1511
              ||  || | |||||| |  | |  ||  | ||||||| |  |||| ||      | ||| |
Db      10474 AAATAT-ATAAATATAAATATATAAATATATAAATATATAAATATATAAATATATAAATA 10416

Qy       1512 ATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCC 1571
                | | ||   ||| ||||||   ||       | |  |     ||   ||  ||||||
Db      10415 TAAATATATAAATATATAAATATACATTTGTATATTTATATATATCTATATATAAATATA 10356

Qy       1572 TATATGTATATTTGAAATTAAAAACAGATA 1601
              ||| |  |||| |  |  || | | | |||
Db      10355 TATTTACATATATCTATATATATAAATATA 10326
```

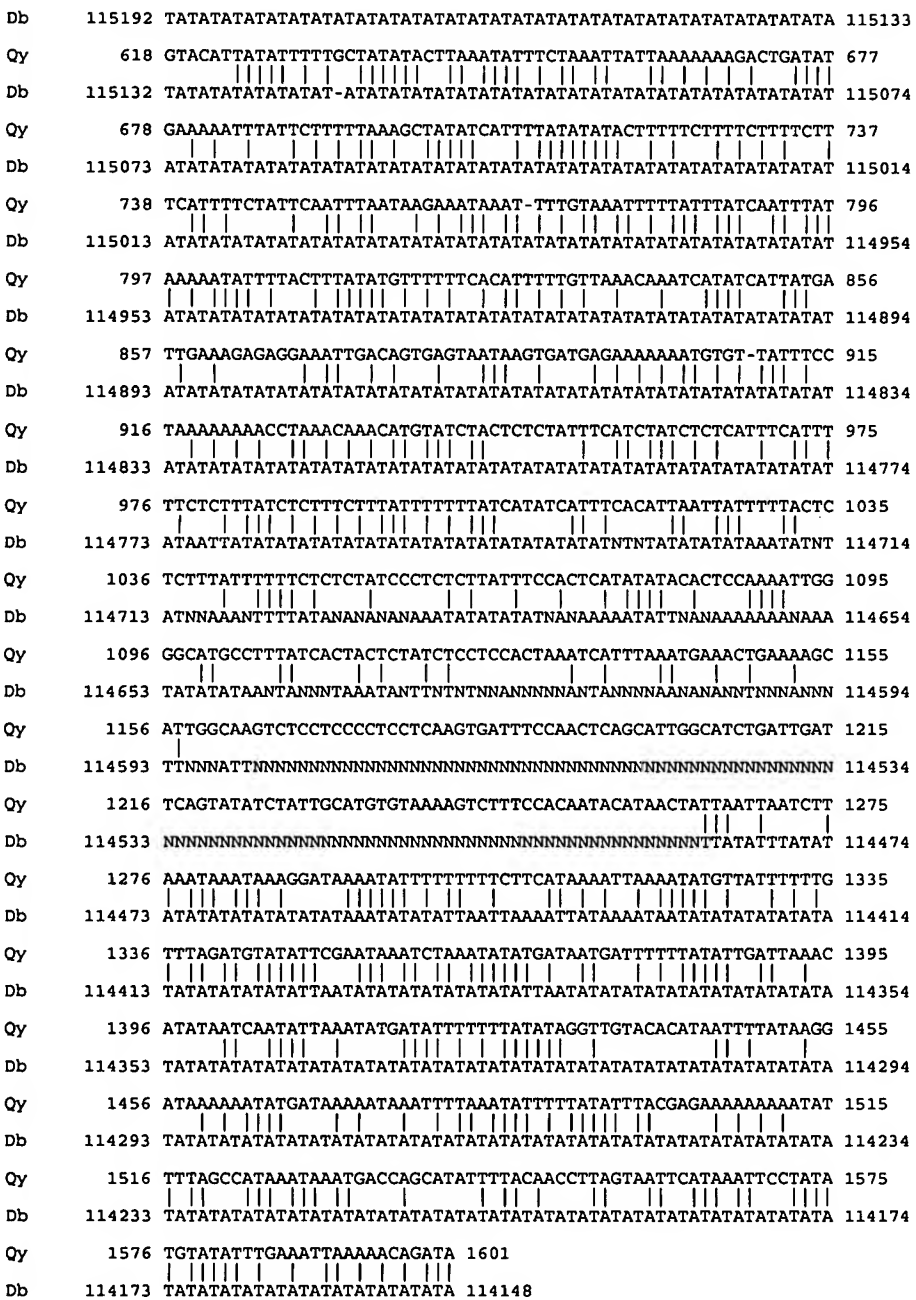
```
RESULT 9
AE014849
LOCUS       AE014849              250421 bp    DNA     linear   INV 11-FEB-2003
DEFINITION  Plasmodium falciparum 3D7 chromosome 12, section 6 of 9 of the
            complete sequence.
ACCESSION   AE014849 AE014188
VERSION     AE014849.1  GI:23496828
KEYWORDS    .
SOURCE      Plasmodium falciparum 3D7
  ORGANISM  Plasmodium falciparum 3D7
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 250421)
  AUTHORS   Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W.,
            Carlton,J.M., Pain,A., Nelson,K.E., Bowman,S., Paulsen,I.T.,
            James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A.,
            Kyes,S., Chan,M.S., Nene,V., Shallom,S.J., Suh,B., Peterson,J.,
            Angiuoli,S., Pertea,M., Allen,J., Selengut,J., Haft,D.,
            Mather,M.W., Vaidya,A.B., Martin,D.M., Fairlamb,A.H.,
            Fraunholz,M.J., Roos,D.S., Ralph,S.A., McFadden,G.I.,
            Cummings,L.M., Subramanian,G.M., Mungall,C., Venter,J.C.,
            Carucci,D.J., Hoffman,S.L., Newbold,C., Davis,R.W., Fraser,C.M. and
            Barrell,B.
  TITLE     Genome sequence of the human malaria parasite Plasmodium falciparum
  JOURNAL   Nature 419 (6906), 498-511 (2002)
  MEDLINE   22255705
   PUBMED   12368864
REFERENCE   2  (bases 1 to 250421)
  AUTHORS   Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
            Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-SEP-2002) Stanford Genome Technology Center, Stanford
            University, 855 California Avenue, Palo Alto, CA 94304, USA
REFERENCE   3  (bases 1 to 250421)
  AUTHORS   Hyman,R.W., Fung,E., Conway,A., Kurdi,O., Mao,J., Miranda,M.,
            Nakao,B., Rowley,D., Tamaki,T., Wang,F. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-JAN-2003) Stanford Genome Technology Center, Stanford
            University, 855 California Avenue, Palo Alto, CA 94304, USA
FEATURES             Location/Qualifiers
     source          1. .250421
                     /organism="Plasmodium falciparum 3D7"
                     /mol_type="genomic DNA"
                     /isolate="3D7"
                     /db_xref="taxon:36329"
                     /chromosome="12"
     repeat_region   188. .292
                     /rpt_type=tandem
     repeat_region   240. .289
                     /rpt_type=tandem
     repeat_region   371. .432
                     /rpt_type=tandem
     repeat_region   439. .527
                     /rpt_type=tandem
     repeat_region   922. .951
                     /rpt_type=tandem
     repeat_region   951. .994
                     /rpt_type=tandem
     repeat_region   1255. .1287
                     /rpt_type=tandem
     gene            <1338. .>3806
                     /locus_tag="PFL1480w"
     mRNA            <1338. .>3806
                     /locus_tag="PFL1480w"
     CDS             1338. .3806
                     /locus_tag="PFL1480w"
                     /note="match to ESTs AU087247.1, AU087385.1"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAN36382.1"
                     /db_xref="GI:23496829"
                     /translation="MNELVVFDTYHGRSINDCELDYYSKKLATCSSDNTVKIFDVSLS
                     KEPVCVAELKDHSSAVWKVCWSHPKYGSLLASCSFDKNVIIYKEVNINKYEMIYINNE
                     HVSSVNYIEWSPHEYGLHLGCASSDGTISILSYYMNKGSNEGYWNKYSMKAHLNGVAC
                     ISWEKPYNNILLMNKNLNDNNNNNNNNNNNNNNNNNNNNNDVINSFKLVSGGYDNQ
                     VIIWMFDNNTKEFQKIYQMNDKPHKSSIKDVAWKPNLDNSANIIASCSDEKIVILWIE
                     DVTNNVWKNGQIIKVKYNIHKISWSPNGTILAIACSDDNAYLYKENAEGIWEEMCNLT
                     EENNNIQDDNSMNMTNNEQENHNENMLNNDNNNMNYTGNVNVDSASYMHDNTNASNL
                     NMNYPNGSNSLINPNNNLSHQGPYQNKDTFHRSSQGNFAAYNNSTQQQQQYSNVNSTG
                     SQMNNVQAGNTVSNNVISNFANNNTNQMKGPPGQMNEPPGQMNGPPSQMNGPPSQMNG
                     PPSQITVPPTPPAYLTMQNSKKNNTSNLTPNMNNIPPPTAPPLVVNNSHGASGPYAPS
                     NFSQQPHGVSAPFGVPSSSPPAQQQVGGSVTNMSVSSNINNNKSTFGGPPPPLPSTST
                     TPIIPAAGSSNFPKASTSSLQQQVHGPFPNVNMSRSSFAINSQMPPSFSHPNNNNNNN
                     NNNNINNNNNNSGMMPPRPSQYSNNLNLNPLSTSVPPPAFSSDTQGMQNKNSITSNFN
                     NNSNNSFSNFNSHMNQANLPPPNPSVMNPSKPVGSTYGSISGGYNVPHAPNMMNDSNV
                     KPQNQQYGYSNYNMPPRQNTHMNN"
     repeat_region   1878. .1943
                     /rpt_type=tandem
     repeat_region   1884. .1942
                     /rpt_type=tandem
     repeat_region   2765. .2875
                     /rpt_type=tandem
     repeat_region   3366. .3417
                     /rpt_type=tandem
     repeat_region   3949. .3978
                     /rpt_type=tandem
     repeat_region   4114. .4144
                     /rpt_type=tandem
     repeat_region   4164. .4197
                     /rpt_type=tandem
     repeat_region   4427. .4464
                     /rpt_type=tandem
     repeat_region   4531. .4558
                     /rpt_type=tandem
     repeat_region   4560. .4589
                     /rpt_type=tandem
     repeat_region   4711. .4740
                     /rpt_type=tandem
     repeat_region   4770. .4800
                     /rpt_type=tandem
     repeat_region   4875. .4920
                     /rpt_type=tandem
     repeat_region   5240. .5289
                     /rpt_type=tandem
     gene            <5379. .>7311
                     /locus_tag="PFL1485w"
     mRNA            join(<5379. .5567,5647. .>7311)
                     /locus_tag="PFL1485w"
     CDS             join(5379. .5567,5647. .7311)
                     /locus_tag="PFL1485w"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAN36383.1"
                     /db_xref="GI:23496830"
                     /translation="MSQDNSDININDEKNEYNDEDIKEYYNEFKVNMMEVKCEGAYRT
                     FKKKKKKKKIYIYIYIYIKKREHVYLACISSLEKEVCRIHTYLNEWRSMHSYEDEEKK
                     SSTGKSTVNINSLRNILIDPCINLEIKELRHKIYEITKKCNIAEEKLQGCNFDAQSIA
                     GQRLINKCKKLQEENYELGKTIEENNLQSLSIQITKLKQHIVFYKNELRVLKELNADI
                     DEDNELLSQQLSEITKKYTLLSEEKMELQKNNDELNNYIKTLKLKYNESYNDIQKDYH
                     KNSYSNINISKKKRNHNIFSYQNDNQENNLHYENVIKGVDNLNNEYNFKYYDYDNNSK
                     HEYSNKYGPHSKYNKHHSNKYDKYDQHNNMYDYNNNNNNSNAYEHNNKYDQYNNTYDY
                     NNKYEPQRYEEYRHNKDERYNNEQDPKTEPQSNINYDKDNNSGRKRNSIDSSRDRKYR
                     KVDKQAYKDKKYEKYKLKDKEKHRKEKTRDKDRERDKERDKDRDKDRDKDRDKDRDKD
                     RDKDRDKERNKDKERNKEKDRHKDRNKDKNKNKYKTKSKMKSKRGGNYDDEEDYDSST
                     YSDRKKKVKNYDDEEYEDFSRGGNKYDNTVTSKNDDMKDATSEHDRYVYMH"
     repeat_region   5514. .5538
                     /rpt_type=tandem
     repeat_region   5538. .5564
                     /rpt_type=tandem
     repeat_region   6900. .7007
                     /rpt_type=tandem
     repeat_region   6910. .6995
                     /rpt_type=tandem
     repeat_region   6914. .7066
                     /rpt_type=tandem
     repeat_region   7316. .7389
                     /rpt_type=tandem
```

```
     repeat_region   7333. .7386
                     /rpt_type=tandem
     repeat_region   7438. .7489
                     /rpt_type=tandem
     repeat_region   7595. .7643
                     /rpt_type=tandem
     repeat_region   7801. .7840
                     /rpt_type=tandem
     repeat_region   7830. .7863
                     /rpt_type=tandem
     repeat_region   7976. .8015
                     /rpt_type=tandem
     repeat_region   8179. .8230
                     /rpt_type=tandem
     repeat_region   8230. .8255
                     /rpt_type=tandem
     repeat_region   8392. .8485
                     /rpt_type=tandem
     repeat_region   8588. .8639
                     /rpt_type=tandem
     repeat_region   9033. .9059
                     /rpt_type=tandem
     gene            <9086. .>11884
                     /locus_tag="PFL1490w"
     mRNA            <9086. .>11884
                     /locus_tag="PFL1490w"
     CDS             9086. .11884
                     /locus_tag="PFL1490w"
                     /note="HMMPfam hit to PF01163, RIO1/ZK632.3/MJ0444 family,
                     score 2.2e-55"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAN36384.1"
                     /db_xref="GI:23496831"
                     /translation="MEQYENYEEIKVLSNDVQTSIIQGEKLTNKFRNRGLTRDKRATI
                     NSVLDNRTLIILRKLKENVYNDIYGVISSGKEAYVFNAIKYINDEELNSLKNIFIKHL
                     EKYKNNINIKINNNENMIYNQMENIHDQCSDVHNYMSQDSERDINNYEQSEDTHFYNN
                     DKYYNGFFNKNKNNKGDQSGDDTNNDIYNDIYNDIYNDIYDDLYNDIYNDIYNDDDEK
                     QNFSKREDTVLEEKNKSNNTYEQTFDIIKEMNDISFYDNKKSQISHIINILDKTKEKK
                     GIALATKVYNTSILVFKKRSQYIEGEFRFRNAYTKNTNPRKMVKQWAEKEFRNLRRIL
                     ICGLRCPYPLVLRSNVIVMSMIGYIDNACPKMKDLNFDILKWKELYIECICILRFLFF
                     NCKLVHADFSEYNLLYFCNHIYIIDVSQSMEHDHPYSLEFLKRDCLNITNFFKKKIGT
                     IVQQTQPTNLYTQRNEQNVENKTNVDINRVQTNYTNEHVEGYNVEGYNVEGHN
                     VEGYNVEGYNVEGHNVEGHNVEGHNVEGHNVEGYDDTNTIDDKTSHLLPQKDFSKLLD
                     VPDAGFYENVEILPLKVLFDFIVSSSLPDDIVYFIENDKKKISLNPYEIIYLQIFGLI
                     KNTTPIPKLNLKKIQKNRIYFEKLKRATCYYVTKFTNEKKKKLEKYSKKHSQKYAHKD
                     EVEEQVFLSSWIPSYLNEIKDIRTIEKDLKLLKKGKSIVNNFISNQNNDHKYQNNTKE
                     EDPKNKKKEQNKEQNKEQNKEQKEIYNKNEVHIDSHIFYEERNQMDVNKINDTLHISN
                     EHSFEENIKQYDHEYKETHQEEIIIEQDKNEENEIMSNSNTLLNIIQTNSCDHLSNSV
                     DSDTSNIDDESDSCEQNEDEENYSSVHEQEEVKFKGIIPDGITRKEWCKLVKEQNREK
                     RKHKIPKYQKKKKKKKLT"
     repeat_region   9636. .9727
                     /rpt_type=tandem
     repeat_region   10548. .10711
                     /rpt_type=tandem
     repeat_region   11322. .11373
```

```
  Query Match             8.0%;  Score 160.4;  DB 3;  Length 250421;
  Best Local Similarity   46.6%;  Pred. No. 4.8e-10;
  Matches  736;  Conservative    0;  Mismatches  816;  Indels   26;  Gaps    6;

Qy         95 AGAAAAAAAAAAAAAAGAAAAGGATCATTTTGAAAGGATATTTTTCGCTCCTATTCAAATA 154
              | |  || ||  ||   |||  || | ||   || |||| ||||       |  |  |||
Db      25178 ATATTAATAATTAATTTAAAATATAAATTAAGAAAGATAATTTTATACTTTATTTAATT 25237

Qy        155 CTGTATTTTTACCAAAAAAACTGTATTTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTC 214
              |||  |       ||| || | ||| ||   ||    |   ||  ||| || ||    ||
Db      25238 AAATATATAGTAATAAATAATTTTATGTTATTTATTATAATAATATTTATTATTTTATTT 25297

Qy        215 GACTCTCATGATTTCCTTCATATGCCAATCACTCTATTTAT----AAATGGCATAAGGTA 270
              | |    | ||         ||   |   | | |||||||    || |   |||
Db      25298 TATTTATTTAATAAATAATTAATTTTATAAAATATATTTATTTTAAATTAAAATATAAAC 25357
```

```
Qy        271 GTGTGAACAATTGCAAAGCTTGTCATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCA 330
              | | |  ||||    |  |    ||  |    ||  ||| ||  || |||| | |||||
Db      25358 ATATAATTAATTAATTAAATATATATATATTTTTTTTAATATAATAAATAATATATTTCA 25417

Qy        331 TGCCTT--TCAAAATTATCTGCACCCCCTAGCTATTAATCTAACATCTAAGTAAGGCTAG 388
                 ||  | |||||  |      |||   ||   |||||  |||     ||  |||    |
Db      25418 CATTTTAATTAAAATAAAAATAACCATTTATTAATTAACTTAATTAATATATAAAATAAA 25477

Qy        389 TGAATTTTTTCGAATAGTCATGCAGTGCATTAATTTCCCCGTGACTATTTTGGCTTTGAC 448
              |||||  | |  || | |          || ||         | |  |||     | |
Db      25478 ATAATTTAATTGTGTAATTAAATTAAATATAAAACATTTATTAATAATTAATTATATATA 25537

Qy        449 TCCAACACTGGCCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTC 508
                    |    | || ||  |   |  ||  |  |    |||||  |   ||||
Db      25538 ATATTATATATTATCTTAAATTAATTAATTTTTTTAATTATTTTAATATAATAATTATTT 25597

Qy        509 TTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTT 568
              | | ||  |||  |||   || ||   |  | ||           |||||  || ||
Db      25598 ATTAATATTAATTATGTATATTTTATTTAATTGTTTA----------ATATTTATTATTA 25647

Qy        569 CTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATAT 628
              | |||| ||   || |   || || | ||||||     |||| |||     ||  |||
Db      25648 TTTTATTTAATATATTATTAATTTAATTAATTATTATATTATATTTAATTATTATATTAT 25707

Qy        629 TTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTAT 688
              || |  |||||  |||| |   |   ||  |  |  |     |||| |  | || ||
Db      25708 TTATAATATATTATTAATTTAATTATTGTTTATTTATTATATTATATATTATTATTTAAT 25767

Qy        689 TCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTAT 748
              | || ||  |  |||   ||| ||||      ||| | |  | |  ||| ||| | |||
Db      25768 TATTATTTTATTTATTATATTATATATTATTAATTTATATAATATAATTTAATTATATAT 25827

Qy        749 TCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTA 808
              | || | |  |   ||| |||     |||||   | |||   | | |||    ||| ||
Db      25828 ATATTTAATTTATTTATATATT-----AATTTAATTATATATTTATTTAATTTAATTATA 25882

Qy        809 CTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGG 868
              || ||| | ||     ||||| |||   |  | || | |||||  |  |    |
Db      25883 TATTTTATTTATTAATTTATTTTATTTATTAATTTAATTTAATTATATATATATTTAATT 25942

Qy        869 AAATTGACAGTGAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTA 928
              ||||      |   |  | || | | | | || |  | |       | | | |  ||
Db      25943 TAATTATATATATATTTAATTTAATTATATATATATTTAATTTAATTGTATATATATTTA 26002

Qy        929 AACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTC 988
              |  |          |  | |   ||   | ||| | | |||  || ||| | |||| |
Db      26003 ATTTATTTATATATTTATTTAATTTATTTATATATTTAATTATATATTTATTTTTATTTT 26062

Qy        989 TTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTC 1048
              ||       |  ||||  |  | | ||   | |  |  ||| |   ||  ||| |  |
Db      26063 TTATAACTATAAATTATTAATTTAATTATTAATTTTATTTTATTTACTAAATATATAATT 26122

Qy       1049 TCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTAT 1108
              | |||   | |  |  ||  | | || || | | | ||   |       |   ||
Db      26123 AATTTATATATTATTGTTTTAATTATTTAAATAATTCATTTTATTTAATTAATTTATA 26182

Qy       1109 CACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTC 1168
              | || | |   |  | | ||| | ||   ||| ||      ||  |  |
Db      26183 TATTATTATTAATAATTCTTTAATTTATAATAATTAATTGTTTAATATATTATTATTAT 26242

Qy       1169 CTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTAT 1228
              |   | ||     ||        ||   || | |||  || | ||  |  |   |
Db      26243 TTAATTATTTAAATATTGTTAATTAATTAATTTATTATATTATTATTTAATTAATTTATT 26302

Qy       1229 TGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGG 1288
              | |     | |   |  | | ||       ||||| ||||| |||| ||||
Db      26303 ATATTATTATTTAATTAATTTATATTAATATTTTATTATTTAATTTAATTAAAATTTAT 26362

Qy       1289 ATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTT--GTTTAGATGTAT 1346
              ||  ||  |||| |||  |  ||| |||| ||  || ||| ||   ||     | ||
Db      26363 TTATTATTATTTTATTTTATATTAATATTAGTATTATTTTTTTATTAATTTATATTTTAA 26422

Qy       1347 ATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAAT 1406
```

```
Db      26423 ATTTAATTATATTAATTTATTATATATTATTAATTAATTTATATTAATATTTTATTATTT 26482

Qy       1407 ATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATAT 1466

Db      26483 ATAAATTATAATATTTTTATTTATTTAATATTTAAATTGTATTAATATTTTTCTATTTAT 26542

Qy       1467 GATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAAT---ATTTTAGCCA 1523

Db      26543 TATTTATATTAATATTCATATTTTATTAATTAATGTGTATTATGAATTTAATTATATTAA 26602

Qy       1524 TAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATT 1583

Db      26603 TATTTAAATTTATTATTATTTTTATTTATTAAATAATTAACTTAATTATTTATTAATATA 26662

Qy       1584 TGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGT 1643

Db      26663 TATATATTTAATTTTAATATATTTTAAATAAATTTTATAATGATTTAATTTGATATATTA 26722

Qy       1644 TTTTCATGCAAACAGAAA 1661

Db      26723 TATAATTATATATATATA 26740


RESULT 10
CR626872/c
LOCUS       CR626872              180096 bp    DNA     linear   HTG 27-JUL-2004
DEFINITION  Danio rerio clone DKEY-38L22, *** SEQUENCING IN PROGRESS ***, 12
            unordered pieces.
ACCESSION   CR626872
VERSION     CR626872.1  GI:50724874
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 180096)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK38L22
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 175236 bases at least Q40
            Consensus quality: 176399 bases at least Q30
            Consensus quality: 177283 bases at least Q20
            Insert size: 178996; sum-of-contigs
            Insert size: 184477; 3.4% error; agarose-fp
            Quality coverage: 4.68x in Q20 bases; sum-of-contigs Quality
            coverage: 4.66x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 12 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    19704: contig of 19704 bp in length
            *    19705    19804: gap of 100 bp
            *    19805    31589: contig of 11785 bp in length
            *    31590    31689: gap of 100 bp
            *    31690    38161: contig of 6472 bp in length
            *    38162    38261: gap of 100 bp
            *    38262    60431: contig of 22170 bp in length
            *    60432    60531: gap of 100 bp
            *    60532    85983: contig of 25452 bp in length
            *    85984    86083: gap of 100 bp
            *    86084    91753: contig of 5670 bp in length
            *    91754    91853: gap of 100 bp
            *    91854   114485: contig of 22632 bp in length
            *   114486   114585: gap of 100 bp
            *   114586   125670: contig of 11085 bp in length
            *   125671   125770: gap of 100 bp
            *   125771   132626: contig of 6856 bp in length
            *   132627   132726: gap of 100 bp
            *   132727   142378: contig of 9652 bp in length
            *   142379   142478: gap of 100 bp
            *   142479   175518: contig of 33040 bp in length
            *   175519   175618: gap of 100 bp
            *   175619   180096: contig of 4478 bp in length.
FEATURES             Location/Qualifiers
     source          1. .180096
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-38L22"
                     /clone_lib="DanioKey"
     misc_feature    1. .19704
                     /note="assembly_fragment:00533
                     fragment_chain:1"
     misc_feature    19805. .31589
                     /note="assembly_fragment:00302
                     fragment_chain:1"
     misc_feature    31690. .38161
                     /note="assembly_fragment:00091
                     fragment_chain:1"
     misc_feature    38262. .60431
                     /note="assembly_fragment:00381
                     fragment_chain:1"
     misc_feature    60532. .85983
                     /note="assembly_fragment:00844
                     fragment_chain:1"
     misc_feature    86084. .91753
                     /note="assembly_fragment:00055
                     fragment_chain:1"
     misc_feature    91854. .114485
                     /note="assembly_fragment:00687
                     fragment_chain:1"
     misc_feature    114586. .125670
                     /note="assembly_fragment:00234
                     fragment_chain:1"
     misc_feature    125771. .132626
                     /note="assembly_fragment:00133
                     fragment_chain:1"
     misc_feature    132727. .142378
                     /note="assembly_fragment:00176
                     fragment_chain:1"
     misc_feature    142479. .175518
                     /note="assembly_fragment:01036
                     fragment_chain:1"
     misc_feature    175619. .180096
                     /note="assembly_fragment:00023.0"
ORIGIN

  Query Match             7.9%;  Score 159.8;  DB 2;  Length 180096;
  Best Local Similarity   43.3%;  Pred. No. 6.3e-10;
  Matches  479;  Conservative    0;  Mismatches  624;  Indels    3;  Gaps    3;

Qy        498 TGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTAT 557

Db     115252 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 115193

Qy        558 ATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAA 617
```



RESULT 11
CR392006/c

```
LOCUS       CR392006             178427 bp    DNA     linear   HTG 23-AUG-2004
DEFINITION  Danio rerio clone DKEYP-57F11, *** SEQUENCING IN PROGRESS ***, 9
            unordered pieces.
ACCESSION   CR392006
VERSION     CR392006.3  GI:51510714
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 178427)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-AUG-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 23, 2004 this sequence version replaced gi:46879107.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zKp57F11
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 173955 bases at least Q40
            Consensus quality: 174654 bases at least Q30
            Consensus quality: 175292 bases at least Q20
            Insert size: 177627; sum-of-contigs
            Insert size: 177796; 2.7% error; agarose-fp
            Quality coverage: 7.97x in Q20 bases; sum-of-contigs Quality
            coverage: 8.05x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    11505: contig of 11505 bp in length
            *    11506    11605: gap of 100 bp
            *    11606    24931: contig of 13326 bp in length
            *    24932    25031: gap of 100 bp
            *    25032    31819: contig of 6788 bp in length
            *    31820    31919: gap of 100 bp
            *    31920    50297: contig of 18378 bp in length
            *    50298    50397: gap of 100 bp
            *    50398    83841: contig of 33444 bp in length
            *    83842    83941: gap of 100 bp
            *    83942   122559: contig of 38618 bp in length
            *   122560   122659: gap of 100 bp
            *   122660   166383: contig of 43724 bp in length
            *   166384   166483: gap of 100 bp
            *   166484   175663: contig of 9180 bp in length
            *   175664   175763: gap of 100 bp
            *   175764   178427: contig of 2664 bp in length.
FEATURES             Location/Qualifiers
     source          1. .178427
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEYP-57F11"
                     /clone_lib="DanioKeypilot"
     misc_feature    1. .11505
                     /note="assembly_fragment:00415
                     fragment_chain:1"
     misc_feature    11606. .24931
                     /note="assembly_fragment:00266
```

```
                    fragment_chain:1"
     misc_feature   25032. .31819
                    /note="assembly_fragment:00041
                    fragment_chain:1"
     misc_feature   31920. .50297
                    /note="assembly_fragment:00579
                    fragment_chain:1"
     misc_feature   50398. .83841
                    /note="assembly_fragment:00848
                    fragment_chain:1"
     misc_feature   83942. .122559
                    /note="assembly_fragment:01340
                    fragment_chain:1"
     misc_feature   122660. .166383
                    /note="assembly_fragment:01877
                    fragment_chain:1"
     misc_feature   166484. .175663
                    /note="assembly_fragment:00124
                    fragment_chain:1"
     misc_feature   175764. .178427
                    /note="assembly_fragment:00006"
ORIGIN

  Query Match             7.9%;  Score 159.6;  DB 2;  Length 178427;
  Best Local Similarity   45.5%;  Pred. No. 6.7e-10;
  Matches  506;  Conservative    0;  Mismatches  599;  Indels    7;  Gaps    4;

Qy        498 TGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTAT 557
              | | |||||   ||  || | ||||       | || || |  |||||  ||  |||
Db      12215 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 12156

Qy        558 ATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAA 617
                ||  | | | | |||     ||  | ||| |  | ||   ||| |||  |||| || |
Db      12155 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 12096

Qy        618 GTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATAT 677
                |||||  | |  ||||||  || |||| | || ||   || | | | |   ||||
Db      12095 TATATATATATATAT-ATATATATATATATATATATATATATATATATATATATATAT 12037

Qy        678 GAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTT 737
              | |   |   | | | || |  |||||   | |||||||| | |   | | | |   |
Db      12036 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11977

Qy        738 TCATTTTCTATTCAATTTAATAAGAAATAAAT-TTTGTAAATTTTTATTTATCAATTTAT 796
                || |     |  || ||   | | ||| || | | || || | ||| |||  || |||
Db      11976 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11917

Qy        797 AAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGA 856
              | | |||| |   | |||| | | |  | || | | |   |   |  ||||   |||
Db      11916 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11857

Qy        857 TTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGT-TATTTCC 915
              | |       | ||  | |   |  |||  |    | | | | || | | |||  |
Db      11856 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11797

Qy        916 TAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTT 975
              | | | |  || | | | || ||| ||        |  || ||| | |   |  || |
Db      11796 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11737

Qy        976 TTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTC 1035
              | | | ||| | | | | ||| | | |||      || |     ||  ||| | ||
Db      11736 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 11677

Qy       1036 TCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGG 1095
                |  | | | | | |||   | | |   |   |   |||||||  |     | ||
Db      11676 ATATATATATATATATATATATATATATATATCATATATATATATATATATATATATA 11617

Qy       1096 GGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGC 1155
                ||   | |
Db      11616 TATATATATATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 11557

Qy       1156 ATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGAT 1215
                                                                   | ||

Db      11556 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTTATAATA 11497

Qy       1216 TCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTT 1275
              |     ||| |||   |      | | |  |      | | | ||    || ||||| |
Db      11496 TATAATTATATATATTAATATATATAATAATATATATAAATAAAATATATATTTAATATA 11437

Qy       1276 AAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTG 1335
              ||||||||||   ||   ||  | || |  || |  ||   |||||||  |||| ||
Db      11436 TAATAAATAAAATTTATTTTAAATATTATAATTTAATAATAAAAAATATAATATTATTAA 11377

Qy       1336 TTTAGATGTATATTCGAATAAATCTAAATATATGAT----AATGATTTTTTATATTGATT 1391
              || | ||    |||    |  | |  || ||  ||    ||| || | | || |   ||
Db      11376 TTAAAATTATAATTTTTTTTTTTTTTATTTAAAATTATAAATTATATATAATTTATTTT 11317

Qy       1392 AAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTAT 1451
              | |  |||| | || |  ||| |  ||||||| || | |   || || | |||  |  |
Db      11316 ATATTTATATTTAAAAAAAAAAAAAATATTTTATTTTTTTTTTTATATATATATATAAAA 11257

Qy       1452 AAGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAA 1511
              |    || | | |    ||| || |||   |||| ||  |     ||     |||  ||
Db      11256 ATATTTATATATTTAATTAATAAAAAAAAATAAAAATAATATATATTTTTTTAAATGAAT 11197

Qy       1512 ATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCC 1571
              ||| |  |   ||| ||| ||    |       | || |    ||   ||  ||| ||
Db      11196 ATAATAAATATATATATATATATATATATATATATATATATATATATATATATATATA 11137

Qy       1572 TATATGTATATTTGAAATTAAAAACAGATAAT 1603
              ||||| ||||| |  |  ||  |  || ||||
Db      11136 TATATATATATATATATATATTATTAGTTAAT 11105


RESULT 12
HSJ282H10
LOCUS       HSJ282H10             120029 bp    DNA     linear   HTG 10-JUL-2001
DEFINITION  Homo sapiens chromosome 6 clone RP1-282H10 map p22.1-22.3, 5
            unordered pieces.
ACCESSION   AL132672
VERSION     AL132672.14  GI:14348905
KEYWORDS    HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Garner,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 12, 2001 this sequence version replaced gi:12331282.
            -------------- Genome Center
            Center: Sanger Centre
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: dJ282H10
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Sequencing vector: M13; M77815; 0% of reads
            Sequencing vector: plasmid; L08752; 99% of reads
            Chemistry: Dye-terminator ABI; 1% of reads
            Chemistry: Dye-terminator Big Dye; 98% of reads
            Consensus quality: 118224 bases at least Q40
            Consensus quality: 118449 bases at least Q30
            Consensus quality: 118634 bases at least Q20
            Insert size: 119629; sum-of-contigs
            Insert size: 117431; 9.3% error; agarose-fp
            Quality coverage: 18.41x in Q20 bases; sum-of-contigs Quality
            coverage: 19.36x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
```

```
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    23429: contig of 23429 bp in length
            *    23430    23529: gap of 100 bp
            *    23530    63600: contig of 40071 bp in length
            *    63601    63700: gap of 100 bp
            *    63701   100099: contig of 36399 bp in length
            *   100100   100199: gap of 100 bp
            *   100200   102366: contig of 2167 bp in length
            *   102367   102466: gap of 100 bp
            *   102467   120029: contig of 17563 bp in length.
FEATURES             Location/Qualifiers
     source          1. .120029
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /map="p22.1-22.3"
                     /clone="RP1-282H10"
                     /clone_lib="RPCI-1"
     misc_feature    1. .23429
                     /note="assembly_fragment:05837
                     clone_end:SP6
                     vector_side:left"
     misc_feature    23530. .63600
                     /note="assembly_fragment:00423
                     fragment_chain:1"
     misc_feature    63701. .100099
                     /note="assembly_fragment:02685
                     fragment_chain:1"
     misc_feature    100200. .102366
                     /note="assembly_fragment:01789"
     misc_feature    102467. .120029
                     /note="assembly_fragment:02341
                     clone_end:T7
                     vector_side:right"
ORIGIN

  Query Match             7.9%;  Score 158.6;  DB 2;  Length 120029;
  Best Local Similarity   45.6%;  Pred. No. 9.8e-10;
  Matches  508;  Conservative    0;  Mismatches  594;  Indels   13;  Gaps    4;

Qy        488 AAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGT 547
Db      22834 AAAATATATATATTTTTTATATATATATTTTATATATTTTATTTTATATATATTTTAT 22893

Qy        548 TAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATG 607
Db      22894 TTATTTATATATTTATATATTATATATTTTTATATATAATATATAAAATATATATTTTAT 22953

Qy        608 CATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAA 667
Db      22954 ATATTATATATAAATATATATATTATATTTATATATATATTTATATATTATATATAAT 23013

Qy        668 AGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTT 727
Db      23014 ATATATAATATATATGTATTATATATATTTATATATATAATATATTATATATTATATATA 23073

Qy        728 TTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTA 787
Db      23074 TTATATATTATATATTTTATATATTATATATTATATATTATATATTATATTTTATACA 23133

Qy        788 TCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATA 847
Db      23134 TTATATATTATATATTATATATTTTATATATTATATATTTTATATTTTATATTTTATATA 23193

Qy        848 TCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTG 907
Db      23194 TATTATATATTTTATATATCATATATATTATATATCATATATTTTTATATATTATATATT 23253

Qy        908 TTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCA 967
Db      23254 TTATATATTATATATTATATATATTATATATTTTATATATATTATATATTTTATATATTT 23313

Qy        968 TTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTAT 1027
Db      23314 TATATATTTTATATATATTTTATATATTTTATATATTATATATATTTTATATATTATATA 23373

Qy       1028 TTTTACTCTCTTTATTTTTTCTCTCTATCCC------TCTCTTATTTCCACTCATATATA 1081
Db      23374 TTTTACACATTTTATATATTTTATATATTATATAGGATATATATTTTATATTTTATNNNN 23433

Qy       1082 CACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAA 1141
Db      23434 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 23493

Qy       1142 TGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATT 1201
Db      23494 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTTTATATATTATATATTTTATAT 23553

Qy       1202 GGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAAC 1261
Db      23554 ATTATATATTATATATATTATATATTTTATATATATTATATATTTTATATATTTTATATA 23613

Qy       1262 TATTAATTA-ATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320
Db      23614 TTTTATATATATTTTATATATTTTATATATTATATATATTTTATATATTATATATTTTAT 23673

Qy       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380
Db      23674 ATATTTTATATATTTTATATAT-----TATATATAATATATATTTTATATTTTATATATA 23728

Qy       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440
Db      23729 TTATATATTTTATATTTTTTATATATATTATATATTTTATATTTTATATATATTATATA 23788

Qy       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500
Db      23789 TTATATATATTTTATGTATATTATATATTATATATATTTTATATATTATATATTATATA- 23847

Qy       1501 GAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560
Db      23848 TATTTTATATATTATATATTATATATATTTATATATATTATATAGTATATATATTTTATAT 23907

Qy       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAA 1595
Db      23908 ATACTATATAATATATATTTTATATATATTATATA 23942


RESULT 13
AC091214
LOCUS       AC091214               143331 bp    DNA     linear   PRI 07-MAR-2002
DEFINITION  Homo sapiens 12q BAC RP11-13F9 (Roswell Park Cancer Institute Human
            BAC Library) complete sequence.
ACCESSION   AC091214
VERSION     AC091214.20  GI:14786798
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 143331)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
            Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
```

Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newtson,J., Newtson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shooshtari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Kucherlapati,R., Weinstock,G. and Gibbs,R.

TITLE Direct Submission
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 143331)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 143331)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 4 (bases 1 to 143331)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (28-JUL-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 5 (bases 1 to 143331)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (12-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 6 (bases 1 to 143331)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (27-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 7 (bases 1 to 143331)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (31-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 8 (bases 1 to 143331)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (07-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT On Jul 17, 2001 this sequence version replaced gi:14578078.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE:Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES:This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html.

QUALSTAT-REPORT.

FEATURES Location/Qualifiers
source 1. .143331
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-13F9"
misc_feature complement(1. .2005)
/note="overlaps bases 1. .2005 of clone AC079409"
/function="clone overlap"
repeat_region 511. .703
/rpt_family="LTR16A"
repeat_region complement(825. .1128)
/rpt_family="AluY"
repeat_region complement(3191. .3292)
/rpt_family="L2"
repeat_region 3856. .3887
/rpt_family="AT_rich"
repeat_region complement(4061. .4369)
/rpt_family="LTR33A"
repeat_region 4457. .5350
/rpt_family="L1M1"
repeat_region 5351. .5373
/rpt_family="T-rich"
repeat_region complement(5374. .6088)
/rpt_family="L1PA7"
repeat_region 6109. .6146
/rpt_family="(TTTTA)n"
STS 7538. .7716
/standard_name="61535"
repeat_region 12648. .13327
/rpt_family="L1M1"

```
     repeat_region   13353. .13387
                     /rpt_family="AT_rich"
     repeat_region   13423. .15610
                     /rpt_family="L1M1"
     repeat_region   15611. .15710
                     /rpt_family="FLAM_C"
     repeat_region   15718. .15889
                     /rpt_family="(TA)n"
     repeat_region   15909. .16212
                     /rpt_family="(TA)n"
     repeat_region   complement(16213. .16263)
                     /rpt_family="FLAM"
     misc_feature    16261. .17578
                     /function="unresolved tandem repeat"
     repeat_region   16264. .16399
                     /rpt_family="(TA)n"
     repeat_region   16405. .16446
                     /rpt_family="AT_rich"
     repeat_region   16447. .17315
                     /rpt_family="(TA)n"
     repeat_region   17336. .17504
                     /rpt_family="(TA)n"
     repeat_region   17599. .17636
                     /rpt_family="(TA)n"
     repeat_region   17648. .17764
                     /rpt_family="AluJo/FRAM"
     repeat_region   17774. .17807
                     /rpt_family="AT_rich"
     repeat_region   17811. .18491
                     /rpt_family="L1"
     repeat_region   18494. .18612
                     /rpt_family="AluJb"
     repeat_region   18613. .19807

  Query Match             7.9%;  Score 158.4;  DB 9;  Length 143331;
  Best Local Similarity   49.8%;  Pred. No. 9.8e-10;
  Matches  543;  Conservative    0;  Mismatches  531;  Indels   16;  Gaps    5;

Qy        480 TTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTT 539
              |||  |  |   | |||||| ||| |    | | || ||||||  |     | || |  |
Db      16468 TTATTTATATTTATATATTATATATTTATATTAATATAAATATATTTATATATATAAATA 16527

Qy        540 CATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTAT 599
               || ||  | |    |||  ||  | ||| | |||   | ||  | ||| ||  |  |||
Db      16528 TATTTATATTACATATATTATATATATTTATATATATAAATATATTTATATTACATATAT 16587

Qy        600 TATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTA 659
              ||| |||   | | || |  |   ||| | | |  ||||| |  |    ||   |  |
Db      16588 TATATATATTTTTATATATAAATATATTTATATTACATATATTATATCTATTTATATATT 16647

Qy        660 TTAAAAAAAGACTGATATGAAAAA----TTTATTCTTTTTAAAGCTATATCATTTTATAT 715
              |||   | |      |||| | | |    | |||   || || |  | |   || |  |||
Db      16648 TTATGTATATTTATATATTATATATATTTGTATATATTATATATATTTTATATATATTAT 16707

Qy        716 ATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTA 775
              |||  | || | |   | | | |  |||| |||  | ||||   |    || || |
Db      16708 ATATATATTTTATATGTATTTATATTTTTATATATATTTTATATATCTTTATATATATCT 16767

Qy        776 AATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTT 835
              || | |||||||  || | | | |   || ||  |||||| ||| |  | | | |   |
Db      16768 TATGTATATTTATATATATTTTATATATTTATATATTATATATTTGTATATAATATATAT 16827

Qy        836 AAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGA 895
               | | || || | ||     ||| | | |||  | ||  |   | |   ||     ||
Db      16828 TTATATAT-ATTTTATATATATTTATATAGATTATATATATTTTTATATATTTTATATAT 16886

Qy        896 GAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTC 955
               | | | | | |   ||  || | |     || |   | || |      | | ||| |
Db      16887 TTATATATTATATATATTTATACATATTTTATACATTCATATATTATATATATTTATATA 16946

Qy        956 ATCTATCTCTCATTTCATTTT---TCTCTTTATCTCTTTCTTTATTTTTTTATCATATCA 1012
                | | | |  ||  || ||   | | | ||| | ||| | |||||| | |||  |||
Db      16947 TATTTTATATATTTATATATTATATATTTATATATATTTATATATTTATATATTGTATAT 17006

Qy       1013 TTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCAC 1072
              || | ||  ||  || | |  | | | |||| | | | | |||   | | |||| |
Db      17007 ATTTATATATATATTTATATATTTATATATTGTATATATTTATATATATTTTATATATTG 17066

Qy       1073 TCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAA 1132
               ||||||     |    | |      ||   || ||   || | ||| | ||   || |
Db      17067 ATATATATTATATATTTATATATTTTATATATTGATATATATTATATATATTCATATATA 17126

Qy       1133 TCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATT----- 1187
               | | || ||  |  |  |    ||  | | |  |      |  |  |||
Db      17127 TTTTATATATTTATATATATTATATATTTATATATATTTTATATATTTATATATTATATA 17186

Qy       1188 --TCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCT 1245
                 |  |  |     ||       | ||  |||    || | |||   || | ||    | |
Db      17187 TATTTATATATATTTTATATATTTATATATTATACATTTATATATTATATATATTTATAT 17246

Qy       1246 TTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTT 1305
               |     |||  ||  |||||  || |  || || |    |    || |   | | | | |
Db      17247 ATTATATATATTTATATATTATATATTTATATATTATATATATTTATATATTATATATAT 17306

Qy       1306 CTTCATAAAATTAAAATATGT-TATTTTTTGTTTAGATGTATATTCGAATAAATCTAAAT 1364
               |  ||| | || | |||| | ||| |||| | ||  | ||||||   ||| || |  ||
Db      17307 TTATATATAGTTTATATATTTATATATTTTATATATTTATATATTTTTATATATATTTAT 17366

Qy       1365 ATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTT 1424
               ||||  | ||   |||||||||||| ||| | || ||   | |  || || |     | ||
Db      17367 ATATTTTTATATATTTTTATATTGTTTATATATTTATGTATTTATATATTTATATATATT 17426

Qy       1425 TTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAA 1484
               |||||||  |      | ||| || ||||       ||| |||| ||| ||||| || || |
Db      17427 TTATATATTTATATTATATATTTATATATTATATAAATATATAATATAAATATATGTTTA 17486

Qy       1485 ATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATT 1544
                | | |||  |    |  | | | || || |      | | || |  |     |||
Db      17487 TATTATATATTATATATAATACATATAAAATATATATTTTCAGTAGAGTATACATATAAG 17546

Qy       1545 TTACAACCTT 1554
                ||  |   ||
Db      17547 ATATACATTT 17556
```

```
RESULT 14
CR759967/c
LOCUS       CR759967              226108 bp    DNA     linear   HTG 11-SEP-2004
DEFINITION  Danio rerio clone DKEY-94P13, *** SEQUENCING IN PROGRESS ***, 6
            unordered pieces.
ACCESSION   CR759967
VERSION     CR759967.2  GI:51988233
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
            Cypriniformes; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 226108)
  AUTHORS   Sims,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-2004) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 11, 2004 this sequence version replaced gi:51965792.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: zfish-help@sanger.ac.uk
            -------------- Project Information
            Center project name: zK94P13
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 223262 bases at least Q40
```

```
            Consensus quality: 223728 bases at least Q30
            Consensus quality: 224330 bases at least Q20
            Insert size: 225608; sum-of-contigs
            Insert size: 219705; 1.4% error; agarose-fp
            Quality coverage: 11.47x in Q20 bases; sum-of-contigs Quality
            coverage: 11.83x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     5476: contig of 5476 bp in length
            *     5477     5576: gap of 100 bp
            *     5577    34821: contig of 29245 bp in length
            *    34822    34921: gap of 100 bp
            *    34922    59876: contig of 24955 bp in length
            *    59877    59976: gap of 100 bp
            *    59977   138375: contig of 78399 bp in length
            *   138376   138475: gap of 100 bp
            *   138476   176066: contig of 37591 bp in length
            *   176067   176166: gap of 100 bp
            *   176167   226108: contig of 49942 bp in length.
FEATURES             Location/Qualifiers
     source          1. .226108
                     /organism="Danio rerio"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7955"
                     /clone="DKEY-94P13"
                     /clone_lib="DanioKey"
     misc_feature    1. .5476
                     /note="assembly_fragment:00022
                     fragment_chain:1"
     misc_feature    5577. .34821
                     /note="assembly_fragment:00494
                     fragment_chain:1"
     misc_feature    34922. .59876
                     /note="assembly_fragment:00083
                     fragment_chain:2"
     misc_feature    59977. .138375
                     /note="assembly_fragment:02246
                     fragment_chain:2"
     misc_feature    138476. .176066
                     /note="assembly_fragment:00979"
     misc_feature    176167. .226108
                     /note="assembly_fragment:01544.0"
ORIGIN

  Query Match             7.9%;  Score 158.4;  DB 2;  Length 226108;
  Best Local Similarity   43.3%;  Pred. No. 8.4e-10;
  Matches  479;  Conservative    0;  Mismatches  623;  Indels    4;  Gaps    4;

Qy        498 TGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTAT 557
              | | |||||   ||  || | ||||       | || || |  |||||  ||  ||| .
Db      60508 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 60449

Qy        558 ATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAA 617
                ||  | | | | |||     ||  | ||| |  | ||   ||| |||  |||| || |
Db      60448 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 60389

Qy        618 GTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATAT 677
                |||||  | |  ||||||  || |||| | || ||   || | | |   ||||
Db      60388 TATATATATATATAT-ATATATATATATATATATATATATATATATATATATATATAT 60330

Qy        678 GAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTT 737
              | |  |   | | | || |  |||||   | |||||||||  | |   | | | |   |
Db      60329 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 60270

Qy        738 TCATTTTCTATTCAATTTAATAAGAAATAAAT-TTTGTAAATTTTTATTTATCAATTTAT 796
                || |     |  || ||   | | ||| || | | || || | ||| |||  || |||
```

```
Db      60269 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 60210

Qy        797 AAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGA 856
              | | |||| |   | |||||  | | |  | || | | |  |   |   ||||   |||
Db      60209 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 60150

Qy        857 TTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAAATGTGT-TATTTCC 915
                | |        | ||  | |   |   |||  |    | | | | || | | ||| |
Db      60149 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 60090

Qy        916 TAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTT 975
                | | | |  || | | | || ||| ||         |  || ||| | |   |  || |
Db      60089 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 60030

Qy        976 TTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTC 1035
                | | | ||| | | | | ||| | | |||       |  | ||  ||
Db      60029 ATATATATATATATATATATATATATATATATATATAATATATATATATATATAANNNNNNN 59970

Qy       1036 TCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGG 1095

Db      59969 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 59910

Qy       1096 GGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGC 1155
                                                  |   | ||||| || ||     ||
Db      59909 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTTNTTATTTATATAAANNNNAATTT 59850

Qy       1156 ATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGAT 1215
              |||      |  |          |  ||   | |   |       | ||   |
Db      59849 ATTNTNTNNTNNNNTTNTNANNNTNTTTTTTATTTAATATAAAAAAAATNTTNATNTTT 59790

Qy       1216 TCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTT 1275
              |    ||||  |    |  |     | |          |   |   |       |
Db      59789 TATTAATATTAAAATTANNTNNNNNANTNNNNTTTTTTTTNNTTTTTTNNNNNNNTNNN 59730

Qy       1276 AAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTG 1335
              ||| ||  |||  | ||| | ||  | ||  |   ||||||||||      || || ||
Db      59729 AAA-AATAAAAAAAAAAAANTTTTAATATTNTTAATTAAAATTAAANNTATTTTTTNTTNN 59671

Qy       1336 TTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAAC 1395
              | || ||  | |     ||| |  | | | | |  |  | |||| || |||  |   |
Db      59670 TNTATATNNAAAAAAATATATANATTATTTTTTTTTTTTTTATTTNTTTTATNTAAATATA 59611

Qy       1396 ATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGG 1455
                | |||  ||||  |    |||| | | ||||||  ||          | ||||
Db      59610 TAAAAATATATATATATATATATATATATATATATATNTTTATATATATTATATATATAT 59551

Qy       1456 ATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAATAT 1515
              ||| | |     ||||| ||| || |   |||| | ||||| ||      | | | |
Db      59550 ATATATATATATATAAATATATATATATTATATATATATATATATATATATATATATA 59491

Qy       1516 TTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATA 1575
              | ||   ||| ||| ||     |     | || |   ||   ||  ||| ||   ||||
Db      59490 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATA 59431

Qy       1576 TGTATATTTGAAATTAAAAACAGATA 1601
              | ||||| |  |  || | | | |||
Db      59430 TATATATATATATATATATATATA 59405


RESULT 15
AC084128
LOCUS       AC084128              129240 bp    DNA     linear   PRI 30-JUL-2002
DEFINITION  Homo sapiens chromosome 8, clone CTD-3118D11, complete sequence.
ACCESSION   AC084128
VERSION     AC084128.6  GI:22004605
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 129240)
  AUTHORS   Birren,B., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone CTD-3118D11
```

```
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 129240)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
            DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
            FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
            Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
            Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G.,
            Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
            O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
            Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
            Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
            Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
            Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-OCT-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 129240)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
            Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
            Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,
            Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
            Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,
            Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
            Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
            Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
            Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   4  (bases 1 to 129240)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
            Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B.,
            Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
            Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
            Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
            McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
            Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
            O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
            Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
            Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
            Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
            Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
            Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jul 30, 2002 this sequence version replaced gi:21306731.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L11383
                Center clone name: 3118_D_11
            --------------.
FEATURES             Location/Qualifiers
     source          1. .129240
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="CTD-3118D11"
                     /clone_lib="CITD2 Human BAC"
     repeat_region   1. .224
                     /rpt_family="AluSg/x"
     repeat_region   225. .403
                     /rpt_family="AluJb"
     repeat_region   669. .855
                     /rpt_family="L1MC3"
     repeat_region   1519. .1600
                     /rpt_family="AT_rich"
     repeat_region   complement(1619. .1895)
                     /rpt_family="AluSc"
     repeat_region   complement(1897. .2197)
                     /rpt_family="AluSq"
     repeat_region   2310. .2593
                     /rpt_family="Charlie2"
     repeat_region   complement(2598. .3451)
                     /rpt_family="L1MA4A"
     unsure          2966. .2969
                     /note="single clone coverage"
     repeat_region   3483. .3713
                     /rpt_family="L1MDa"
     repeat_region   complement(3714. .4017)
                     /rpt_family="AluSg"
     unsure          3739. .3748
                     /note="<30 qual SNGL region"
     unsure          3785. .3811
                     /note="<30 qual SNGL region"
     repeat_region   4018. .4255
                     /rpt_family="L1MDa"
     repeat_region   4267. .4982
                     /rpt_family="L1MDa"
     repeat_region   4974. .5464
                     /rpt_family="L1MD2"
     repeat_region   5827. .6015
                     /rpt_family="MIR"
     repeat_region   8136. .8345
                     /rpt_family="MIR"
     repeat_region   8552. .8737
                     /rpt_family="L2"
     repeat_region   9347. .9644
                     /rpt_family="AluSx"
     repeat_region   complement(10188. .10324)
                     /rpt_family="MIR3"
     repeat_region   10770. .11002
                     /rpt_family="MIR"
     repeat_region   complement(11116. .11638)
                     /rpt_family="MER1A"
     repeat_region   12153. .12222
                     /rpt_family="L2"
     repeat_region   complement(12668. .12878)
                     /rpt_family="MIR"
     repeat_region   complement(13676. .14221)
```

```
                     /rpt_family="MER41A"
     repeat_region   complement(15056. .15245)
                     /rpt_family="AluJb"
     repeat_region   15272. .15714
                     /rpt_family="AluSx"
     repeat_region   complement(16162. .16386)
                     /rpt_family="MIR"
     repeat_region   16712. .17008
                     /rpt_family="AluJo"
     repeat_region   complement(18862. .18963)
                     /rpt_family="MER5A"
     repeat_region   complement(19028. .19089)
                     /rpt_family="L2"
     repeat_region   19090. .19375
                     /rpt_family="AluJb"
     repeat_region   complement(19376. .19430)
                     /rpt_family="L2"
     repeat_region   complement(19480. .19643)
                     /rpt_family="L2"
     repeat_region   complement(19744. .19829)
                     /rpt_family="MIR3"
     repeat_region   complement(19976. .20109)
                     /rpt_family="MIR"
     repeat_region   22048. .22074
                     /rpt_family="(T)n"
     repeat_region   complement(22240. .22864)
                     /rpt_family="L1PA7"
     repeat_region   complement(23650. .23959)
                     /rpt_family="AluJb"
     repeat_region   complement(24055. .24374)
                     /rpt_family="AluSg1"
     repeat_region   24894. .24922
                     /rpt_family="AT_rich"
     repeat_region   complement(24934. .25223)
                     /rpt_family="AluJb"
     repeat_region   26058. .26354
                     /rpt_family="AluSx"
     repeat_region   26936. .27229
                     /rpt_family="AluSx"
     repeat_region   28099. .28243
                     /rpt_family="AluSx"
     repeat_region   29591. .29611

  Query Match             7.9%;  Score 158;  DB 9;  Length 129240;
  Best Local Similarity   49.5%;  Pred. No. 1.1e-09;
  Matches  556;  Conservative    0;  Mismatches  550;  Indels   18;  Gaps    5;

Qy        497 TTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTA 556
              || ||||| |   ||| || | |||||      | || ||  | |||    ||  |||
Db      31982 TTATTTATTTATATAAATACATATATTTATAAATATATATAATTATAATTATATATATAT 32041

Qy        557 TATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAA 616
              |  || || | | | | |   | ||  | |||||  ||||   |||||||   ||| ||
Db      32042 TTATAATTATATATATTTATAAATATATTTATTTATAAATATATATTTATAATTATATAT 32101

Qy        617 AGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATA 676
              | |  ||||||  |    |||||   ||| ||| | ||  |  ||| ||| | |    ||
Db      32102 ATTTAATTATAAATA---TATATTTATAATTATATATATTTATTTATAAATATATATTTA 32158

Qy        677 TGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCT 736
              | |  |  |   | | |||| |  |||||  || ||  ||||  | ||  ||| |
Db      32159 TAATTATATATATTTATTTATAAATATATATTTATAATTATATATATTTATTTATAAATA 32218

Qy        737 TTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTAT 796
              |  ||||   |||  || || | |   ||||| |    | || | |||    |||||||
Db      32219 TATATTTATAATTATATATATTTATTTATAAATAT--ATATTTATAATTATATAATTTAT 32276

Qy        797 AAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGA 856
                | || | |||  |   |    | |  | ||| |   || | |||| |||| |||||
Db      32277 TTATATGTATTATATATAAATAATATATATTATATATATATAAATTATATAATTATAT 32336

Qy        857 TTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCT 916
                | |   | |   | ||     | | | |||  | || | | |   || | |     | |
```

```
Db      32337 ATAATTTATATATATTTATATATAAATTATATATAATTATATATTTATATATATAAACAT 32396

Qy        917 AAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTTT 976
              | | |||   || | | || |||    | | ||||| || ||| |   |  | |   |||
Db      32397 ATATAAATTATATATTTATATATATAATTTATATATTT-ATATATATAAATTATATATTT 32455

Qy        977 TCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTCT 1036
                  |   | | | | | | | | |||  |   || |    || ||   | |   | |
Db      32456 ATATATATAATTTATATATTTATATATATATTTATATATATATTTTATATATATATTTAT 32515

Qy       1037 CTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGG 1096
               | ||   | | | | ||    | | |          | ||| ||  |     |  |
Db      32516 ATATAAAATATATATATAAATATATATATAAATATATTATACATTGATATATATTTATAT 32575

Qy       1097 GCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCA 1156
                ||  ||    |  | |    |  |  | | | | | | | |    |  |
Db      32576 ATATAAATTATATATAAATTATATATTTATATATATTAAATATATTTATATATTTATATA 32635

Qy       1157 TTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATT 1216
               |   | | |          | |   ||||  |  ||  ||   ||   ||  |||
Db      32636 TAAATATATTTATATATAAATTATATCTATTTATATATCTATATATAAATTATATATATT 32695

Qy       1217 CAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTA 1276
               |     |    | || | || |    ||| | | |     ||| ||  ||||  ||
Db      32696 TATATATTTATATATATTTATATATAATTTTATATATTTATATACTTATATATAATTATA 32755

Qy       1277 AATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGT 1336
               |||||||  |       |||| | | ||| |  ||||   || |   |  ||| | |
Db      32756 TATAAATATATATATTTATATATATATTTATATATAATTATATATATTTATATATATAAA 32815

Qy       1337 TTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTT----------TATA 1385
               ||| ||  |||||    ||  | || ||| || |||  ||| |||           |||
Db      32816 TTATATAGATATTTATTTATTTATATATAAATTATATAGATATTTATTTATTTATACATA 32875

Qy       1386 TTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAA 1445
               || ||   |  |  | ||  ||||| |||| |||| | |||||||||   | || | |
Db      32876 TTTATATATAAATTTATATATATTATATATTATATATCTTTATATATAATATATATAAAT 32935

Qy       1446 TTTTATA-AGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGA 1504
               || |||| |  ||| | | |   ||| ||||  || |  | ||| | ||||  |
Db      32936 TTATATATATAATATATATTTATATATAAATTTATATATATTATATATATAAATTTATAT 32995

Qy       1505 AAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCAT 1564
               | | |||||||||    ||||   |      |   | | || |   |   || || ||
Db      32996 ATTATATATATTTTATGTATAATATATATTTAAAATTTATATATAGAATATATATTTTAT 33055

Qy       1565 AAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTA 1608
               | |    | ||| |||||||| | | |    | | ||| |   ||
Db      33056 ATACATTTTTATATATATTTTATAATTTTTATATATATTTTATA 33099
```

Search completed: September 3, 2005, 04:25:16
Job time : 6118.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 21:53:41 ; Search time 781.833 Seconds
(without alignments)
15234.088 Million cell updates/sec

Title: US-10-776-889-1
Perfect score: 2012
Sequence: 1 atcttaggcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2012	100.0	2012	13	ADR16015	Adr16015 Soybean a
2	2012	100.0	2012	13	ADR20130	Adr20130 Soybean a
3	1883	93.6	1883	13	ADR16027	Adr16027 Soybean a
4	1719	85.4	1719	13	ADR16028	Adr16028 Soybean a
5	1553	77.2	1553	13	ADR16029	Adr16029 Soybean a
6	1367	67.9	1367	13	ADR16030	Adr16030 Soybean a
7	1160	57.7	1160	13	ADR16031	Adr16031 Soybean a
8	967	48.1	967	13	ADR16032	Adr16032 Soybean a
9	770	38.3	770	13	ADR16033	Adr16033 Soybean a
10	584	29.0	584	13	ADR16034	Adr16034 Soybean a
11	425	21.1	425	13	ADR16035	Adr16035 Soybean a
12	174	8.6	174	13	ADR16036	Adr16036 Soybean a
13	156	7.8	15548	6	ABL34155	Abl34155 Human imm
c 14	150.2	7.5	8056	8	ABZ10246	Abz10246 Haematopo
15	148.6	7.4	8056	8	ABZ10246	Abz10246 Haematopo
16	147.4	7.3	7597	6	ABL33013	Abl33013 Human imm
c 17	142.2	7.1	8056	8	ABZ10100	Abz10100 Haematopo
c 18	136.8	6.8	5286	13	ADS89552	Ads89552 Oligonucl
19	136	6.8	115218	8	ACA64845	Aca64845 Human HNR
20	135.8	6.7	110000	13	ABD32968_6	Continuation (7 of
c 21	135.2	6.7	5286	13	ADS89278	Ads89278 Oligonucl
22	133.8	6.7	5286	13	ADS89278	Ads89278 Oligonucl
23	133.8	6.7	5286	13	ADS89552	Ads89552 Oligonucl
24	133.6	6.6	8056	8	ABZ10100	Abz10100 Haematopo
c 25	133	6.6	158001	12	ADL17884	Adl17884 Human pho
26	131.8	6.6	7814	4	AAS46530	Aas46530 Tumour su
27	128.4	6.4	7498	6	ABL32257	Abl32257 Human imm
c 28	127.8	6.4	99764	13	ABD32992	Abd32992 Human can
29	124.4	6.2	6175	6	ABL33307	Abl33307 Human imm
30	123.8	6.2	6419	6	ABL32267	Abl32267 Human imm
31	123.8	6.2	12237	6	ABL34358	Abl34358 Human imm
c 32	122.4	6.1	1300	12	ADP85917	Adp85917 Synthetic
c 33	122.4	6.1	110000	13	ABD32968_6	Continuation (7 of
c 34	120.6	6.0	1000	12	ADQ62833	Adq62833 Homopoly-
35	120.6	6.0	1000	12	ADQ62832	Adq62832 Homopoly-
36	120	6.0	6109	6	ABL32326	Abl32326 Human imm
37	120	6.0	6109	6	AAS61077	Aas61077 Human gen
38	119	5.9	34548	6	ABL70604	Abl70604 Chemicall
39	118.2	5.9	7442	4	AAS46686	Aas46686 Tumour su
c 40	118	5.9	115218	8	ACA64845	Aca64845 Human HNR
41	116.4	5.8	11222	10	ADB54318	Adb54318 Pretreate
c 42	116.4	5.8	110000	10	ADH10017_0	Adh10017 Human chr
43	116.4	5.8	158001	12	ADL17884	Adl17884 Human pho
44	116	5.8	8222	8	ACF62816	Acf62816 Colon can
45	116	5.8	8222	8	ACF62794	Acf62794 Colon can

ALIGNMENTS

RESULT 1
ADR16015
ID ADR16015 standard; DNA; 2012 BP.
XX
AC ADR16015;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean annexin promoter.
XX
KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS Glycine max.
XX
PN US2004158052-A1.
XX
PD 12-AUG-2004.
XX
PF 11-FEB-2004; 2004US-00776889.
XX
PR 12-FEB-2003; 2003US-0446833P.
XX
PA (KINN/) KINNEY A J.
PA (LIUZ/) LIU Z.
XX
PI Kinney AJ, Liu Z;
XX
DR WPI; 2004-592806/57.
XX
PT New isolated nucleic acids comprising seed-specific soybean annexin and
PT P34 soybean promoters, useful for regulating expression of at least one
PT heterologous nucleic acid fragment in plants.
XX
PS Claim 2; SEQ ID NO 1; 27pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
CC seed-specific soybean annexin promoter or seed-specific soybean P34
CC promoter. The invention also relates to a recombinant expression
CC construct comprising at least one heterologous nucleic acid fragment
CC operably linked to any one of the isolated nucleic acid fragments above,
CC a plant comprising the recombinant expression construct in its genome and
CC a method of regulating expression of at least one heterologous nucleotide
CC sequence in the plant. The plants are dicotyledonous plants, preferably

CC soybean. The method of regulating expression of at least one heterologous
CC nucleotide sequence in a plant comprises transforming a plant cell with
CC the recombinant expression construct, growing fertile mature plants from
CC a transformed plant cell and selecting plants comprising a transformed
CC plant cell expressing the heterologous nucleotide sequence.
CC Alternatively, plants comprising a transformed plant cell expressing the
CC heterologous nucleotide sequence during early seed development are
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC and P34 promoters. The promoters are useful in regulating or promoting
CC expression of at least one heterologous nucleic acid fragment in plants.
CC This sequence represents the soybean annexin promoter of the invention.
XX
SQ Sequence 2012 BP; 695 A; 326 C; 213 G; 778 T; 0 U; 0 Other;

Query Match 100.0%; Score 2012; DB 13; Length 2012;
Best Local Similarity 100.0%; Pred. No. 3.5e-225;
Matches 2012; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60

Qy         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120

Qy        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180

Qy        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240

Qy        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300

Qy        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360

Qy        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420

Qy        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480

Qy        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540

Qy        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600

Qy        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660

Qy        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720

Qy        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780

Qy        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840

Qy        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900

Qy        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960

Qy        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020

Qy       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080

Qy       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140

Qy       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200

Qy       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260

Qy       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320

Qy       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380

Qy       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTTATATAGGTTGTACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTTATATAGGTTGTACA 1440

Qy       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500

Qy       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560

Qy       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620

Qy       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680

Qy       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740

Qy       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800

Qy       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860

Qy       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
```

```
Qy       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980

Qy       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             |||||||||||||||||||||||||||||||||
Db       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012


RESULT 2
ADR20130
ID   ADR20130 standard; DNA; 2012 BP.
XX
AC   ADR20130;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Soybean annexin promoter nucleotide sequence SEQ ID NO:3.
XX
KW   oilseed plant; mature seed; seed fatty acid profile;
KW   polyunsaturated fatty acid; oil; food; food product; beverage;
KW   infant formula; nutritional supplement; pet food; animal feed;
KW   whole bean soy product; aquaculture food product; soybean; annexin;
KW   promoter; ds.
XX
OS   Glycine max.
XX
PN   WO2004071467-A2.
XX
PD   26-AUG-2004.
XX
PF   12-FEB-2004; 2004WO-US005758.
XX
PR   12-FEB-2003; 2003US-0446941P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Kinney AJ,  Cahoon EB,  Damude HG,  Hitz WD,  Kolar CW,  Liu Z;
XX
DR   WPI; 2004-625770/60.
XX
PT   Novel oilseed plant useful for mature seeds in which total seed fatty
PT   acid profile comprises high polyunsaturated fatty acid.
XX
PS   Example 1; SEQ ID NO 3; 132pp; English.
XX
CC   The present invention describes an oilseed plant (I) that produces mature
CC   seeds in which the total seed fatty acid profile comprises at least 1.0%,
CC   5%, 10%, 15%, 20%, 25%, 30%, 40%, 50%, 60% or more of at least one
CC   polyunsaturated fatty acid having 20 or more carbon atoms and 5 or more
CC   carbon-carbon double bonds, and ratio of docosahexanoic
CC   acid:eicosapentaenoic acid (DHA:EPA) is 1:100-860:100 or 1:100-110:100,
CC   where total seed fatty acid profile further comprises less than 2%
CC   arachidonic acid. Also described: (1) seeds (II) obtained from (I); (2)
CC   oil (III) obtained from (II); (3) a recombinant construct (IV) for
CC   altering the total fatty acid profile of mature seeds of an oilseed
CC   plant, comprising two or more promoters, where each promoter is operably
CC   lined to a nucleic acid sequence encoding a polypeptide (e.g., delta-4
CC   desaturase, delta-5 desaturase, delta-6 desaturase, delta-15 desaturase,
CC   delta-17 desaturase, 18-22C elongase or 20-24C elongase) required for
CC   making one or more polyunsaturated fatty acid having 20 or more carbon
CC   atoms and 4 or more carbon-carbon double bonds, and the total fatty acid
CC   profile comprises 2% or more of 1 or more polyunsaturated fatty acid
CC   having 20 or more carbon atoms and 4 or more carbon-carbon double bonds;
CC   (4) an oilseed plant (V) comprising (IV); (5) seeds (VI) obtained from
CC   (V); (6) oil (VII) obtained from (V); (7) a food product or food analogue
CC   (VIII) comprising (II), (III), (VI) or (VII); (8) a beverage comprising
CC   (III) or (VII); (9) infant formula comprising (III) or (VII); (10) a
CC   nutritional supplement comprising (III) or (VII); (11) a pet food
CC   comprising (II), (III), (VI) or (VII); (12) animal feed comprising (III)
CC   or (VII); (13) a whole bean soy product comprising (II) or (VI); and (14)
CC   an aquaculture food product comprising (II), (III), (VI) or (VII). (I) is
CC   useful for producing mature seeds in which the total seed fatty acid
CC   profile comprises polyunsaturated fatty acids. (IV) is useful for
CC   producing an oilseed plant having an altered fatty acid profile which
CC   involves transforming a plant with (IV), growing the transformed plant
CC   and selecting those plants, where the total fatty acid profile comprises
CC   1% or more of one or more polyunsaturated fatty acid having 20 or more
CC   carbon atoms and 5 or more carbon-carbon double bonds. (I) enables the
CC   production of very long chain body unsaturated fatty acids. The present
CC   sequence represents a soybean annexin promoter nucleotide sequence, which
CC   is used in the exemplification of the present invention.
XX
SQ   Sequence 2012 BP; 695 A; 326 C; 213 G; 778 T; 0 U; 0 Other;

  Query Match            100.0%;  Score 2012;  DB 13;  Length 2012;
  Best Local Similarity  100.0%;  Pred. No. 3.5e-225;
  Matches 2012;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCTTAGGCCCTTGATTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAAT 60

Qy         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTTTTCCTTTGAATAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCA 120

Qy        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTGAAAGGATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTAT 180

Qy        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTCCTACACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCC 240

Qy        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATCACTCTATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAA 300

Qy        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGC 360

Qy        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTA 420

Qy        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCAT 480

Qy        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTC 540

Qy        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 ATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATT 600

Qy        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTAT 660

Qy        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACT 720

Qy        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTT 780

Qy        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACA 840

Qy        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900
```

```
Db        841 AATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAA 900

Qy        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 AAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTA 960

Qy        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACAT 1020

Qy       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATAT 1080

Qy       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAA 1140

Qy       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 ATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCAT 1200

Qy       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAA 1260

Qy       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 CTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAA 1320

Qy       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTT 1380

Qy       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 TTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACA 1440

Qy       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 CATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTAC 1500

Qy       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAAT 1560

Qy       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 TCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATC 1620

Qy       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCA 1680

Qy       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 CCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGC 1740

Qy       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 TCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTCACATG 1800

Qy       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTA 1860

Qy       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 ATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAA 1920

Qy       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 GATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAG 1980

Qy       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
              ||||||||||||||||||||||||||||||||
Db       1981 AAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012


RESULT 3
ADR16027
ID   ADR16027 standard; DNA; 1883 BP.
XX
AC   ADR16027;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Soybean annexin promoter truncated sequence #1.
XX
KW   Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS   Glycine max.
XX
PN   US2004158052-A1.
XX
PD   12-AUG-2004.
XX
PF   11-FEB-2004; 2004US-00776889.
XX
PR   12-FEB-2003; 2003US-0446833P.
XX
PA   (KINN/) KINNEY A J.
PA   (LIUZ/) LIU Z.
XX
PI   Kinney AJ,  Liu Z;
XX
DR   WPI; 2004-592806/57.
XX
PT   New isolated nucleic acids comprising seed-specific soybean annexin and
PT   P34 soybean promoters, useful for regulating expression of at least one
PT   heterologous nucleic acid fragment in plants.
XX
PS   Claim 2; SEQ ID NO 13; 27pp; English.
XX
CC   The invention relates to an isolated nucleic acid fragment comprising a
CC   seed-specific soybean annexin promoter or seed-specific soybean P34
CC   promoter. The invention also relates to a recombinant expression
CC   construct comprising at least one heterologous nucleic acid fragment
CC   operably linked to any one of the isolated nucleic acid fragments above,
CC   a plant comprising the recombinant expression construct in its genome and
CC   a method of regulating expression of at least one heterologous nucleotide
CC   sequence in the plant. The plants are dicotyledonous plants, preferably
CC   soybean. The method of regulating expression of at least one heterologous
CC   nucleotide sequence in a plant comprises transforming a plant cell with
CC   the recombinant expression construct, growing fertile mature plants from
CC   a transformed plant cell and selecting plants comprising a transformed
CC   plant cell expressing the heterologous nucleotide sequence.
CC   Alternatively, plants comprising a transformed plant cell expressing the
CC   heterologous nucleotide sequence during early seed development are
CC   selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC   and P34 promoters. The promoters are useful in regulating or promoting
CC   expression of at least one heterologous nucleic acid fragment in plants.
CC   This sequence represents a soybean annexin promoter truncated form of the
CC   invention.
XX
SQ   Sequence 1883 BP; 645 A; 307 C; 193 G; 738 T; 0 U; 0 Other;

  Query Match             93.6%;  Score 1883;  DB 13;  Length 1883;
  Best Local Similarity   100.0%;  Pred. No. 3.2e-210;
  Matches 1883;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        130 GATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTATTTTTCCTAC 189
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GATATTTTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTATTTTTCCTAC 60
```

```
Qy       190 ACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCCAATCACTCT 249
Db        61 ACTCTCAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCCAATCACTCT 120

Qy       250 ATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAAAGCTTGCAA 309
Db       121 ATTTATAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAAAGCTTGCAA 180

Qy       310 TGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGCTATTAATCT 369
Db       181 TGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGCTATTAATCT 240

Qy       370 AACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTAATTTCCCCG 429
Db       241 AACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTAATTTCCCCG 300

Qy       430 TGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCATTACATGAAA 489
Db       301 TGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCATTACATGAAA 360

Qy       490 AGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTA 549
Db       361 AGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTA 420

Qy       550 ATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCA 609
Db       421 ATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCA 480

Qy       610 TATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAG 669
Db       481 TATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAG 540

Qy       670 ACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTT 729
Db       541 ACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTT 600

Qy       730 CTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATC 789
Db       601 CTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATC 660

Qy       790 AATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATC 849
Db       661 AATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATC 720

Qy       850 ATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTT 909
Db       721 ATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTT 780

Qy       910 ATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATT 969
Db       781 ATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATT 840

Qy       970 TCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTT 1029
Db       841 TCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTT 900

Qy      1030 TTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAA 1089
Db       901 TTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAA 960

Qy      1090 AATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTG 1149
Db       961 AATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTG 1020

Qy      1150 AAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTG 1209
Db      1021 AAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTG 1080

Qy      1210 ATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATT 1269
Db      1081 ATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATT 1140

Qy      1270 AATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTAT 1329
Db      1141 AATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTAT 1200

Qy      1330 TTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGA 1389
Db      1201 TTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGA 1260

Qy      1390 TTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTT 1449
Db      1261 TTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTT 1320

Qy      1450 ATAAGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAA 1509
Db      1321 ATAAGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAA 1380

Qy      1510 AAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATT 1569
Db      1381 AAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATT 1440

Qy      1570 CCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCAT 1629
Db      1441 CCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCAT 1500

Qy      1630 CTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATC 1689
Db      1501 CTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATC 1560

Qy      1690 ACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCC 1749
Db      1561 ACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCC 1620

Qy      1750 GTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTTCACATGATTTATTAA 1809
Db      1621 GTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTTCACATGATTTATTAA 1680

Qy      1810 CGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACA 1869
Db      1681 CGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACA 1740

Qy      1870 TGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATT 1929
Db      1741 TGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATT 1800

Qy      1930 TCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAG 1989
Db      1801 TCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAG 1860

Qy      1990 GTTAACTAAGAAGCAATACTTCA 2012
Db      1861 GTTAACTAAGAAGCAATACTTCA 1883
```

RESULT 4
ADR16028
ID ADR16028 standard; DNA; 1719 BP.
XX
AC ADR16028;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean annexin promoter truncated sequence #2.
XX
KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS Glycine max.
XX
PN US2004158052-A1.
XX
PD 12-AUG-2004.
XX
PF 11-FEB-2004; 2004US-00776889.
XX
PR 12-FEB-2003; 2003US-0446833P.

XX
PA (KINN/) KINNEY A J.
PA (LIUZ/) LIU Z.
XX
PI Kinney AJ, Liu Z;
XX
DR WPI; 2004-592806/57.
XX
PT New isolated nucleic acids comprising seed-specific soybean annexin and
PT P34 soybean promoters, useful for regulating expression of at least one
PT heterologous nucleic acid fragment in plants.
XX
PS Claim 2; SEQ ID NO 14; 27pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
CC seed-specific soybean annexin promoter or seed-specific soybean P34
CC promoter. The invention also relates to a recombinant expression
CC construct comprising at least one heterologous nucleic acid fragment
CC operably linked to any one of the isolated nucleic acid fragments above,
CC a plant comprising the recombinant expression construct in its genome and
CC a method of regulating expression of at least one heterologous nucleotide
CC sequence in the plant. The plants are dicotyledonous plants, preferably
CC soybean. The method of regulating expression of at least one heterologous
CC nucleotide sequence in a plant comprises transforming a plant cell with
CC the recombinant expression construct, growing fertile mature plants from
CC a transformed plant cell and selecting plants comprising a transformed
CC plant cell expressing the heterologous nucleotide sequence.
CC Alternatively, plants comprising a transformed plant cell expressing the
CC heterologous nucleotide sequence during early seed development are
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC and P34 promoters. The promoters are useful in regulating or promoting
CC expression of at least one heterologous nucleic acid fragment in plants.
CC This sequence represents a soybean annexin promoter truncated form of the
CC invention.
XX
SQ Sequence 1719 BP; 600 A; 273 C; 173 G; 673 T; 0 U; 0 Other;

Query Match 85.4%; Score 1719; DB 13; Length 1719;
Best Local Similarity 100.0%; Pred. No. 3.4e-191;
Matches 1719; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        294 CATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACC 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CATCAAAAGCTTGCAATGTACAAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACC 60

Qy        354 CCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAG 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CCCTAGCTATTAATCTAACATCTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAG 120

Qy        414 TGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGT 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGCATTAATTTCCCCGTGACTATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGT 180

Qy        474 CCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTA 533
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTA 240

Qy        534 AATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCA 593
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCA 300

Qy        594 AATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCT 653
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCT 360

Qy        654 AAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTAT 713
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 AAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTAT 420

Qy        714 ATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTG 773
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTG 480

Qy        774 TAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTG 833
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTG 540

Qy        834 TTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGAT 893
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGAT 600

Qy        894 GAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATT 953
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATT 660

Qy        954 TCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCAT 1013
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCAT 720

Qy       1014 TTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACT 1073
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACT 780

Qy       1074 CATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAAT 1133
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 CATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAAT 840

Qy       1134 CATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAAC 1193
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAAC 900

Qy       1194 TCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAA 1253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAA 960

Qy       1254 TACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAA 1313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAA 1020

Qy       1314 AATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAA 1373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAA 1080

Qy       1374 TGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGG 1433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 TGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGG 1140

Qy       1434 TTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTA 1493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTA 1200

Qy       1494 TATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCT 1553
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCT 1260

Qy       1554 TAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGA 1613
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 TAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGA 1320

Qy       1614 AGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAAC 1673
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 AGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAAC 1380

Qy       1674 CACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGA 1733
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGA 1440

Qy       1734 AGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTT 1793
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 AGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTT 1500

Qy       1794 TCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGT 1853
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 TCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGT 1560

Qy       1854 ACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAA 1913
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1561 ACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAA 1620

Qy       1914 CTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCT 1973
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1621 CTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCT 1680

Qy       1974 GAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
                |||||||||||||||||||||||||||||||||||||||
Db       1681 GAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 1719


RESULT 5
ADR16029
ID   ADR16029 standard; DNA; 1553 BP.
XX
AC   ADR16029;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Soybean annexin promoter truncated sequence #3.
XX
KW   Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS   Glycine max.
XX
PN   US2004158052-A1.
XX
PD   12-AUG-2004.
XX
PF   11-FEB-2004; 2004US-00776889.
XX
PR   12-FEB-2003; 2003US-0446833P.
XX
PA   (KINN/) KINNEY A J.
PA   (LIUZ/) LIU Z.
XX
PI   Kinney AJ,  Liu Z;
XX
DR   WPI; 2004-592806/57.
XX
PT   New isolated nucleic acids comprising seed-specific soybean annexin and
PT   P34 soybean promoters, useful for regulating expression of at least one
PT   heterologous nucleic acid fragment in plants.
XX
PS   Claim 2; SEQ ID NO 15; 27pp; English.
XX
CC   The invention relates to an isolated nucleic acid fragment comprising a
CC   seed-specific soybean annexin promoter or seed-specific soybean P34
CC   promoter. The invention also relates to a recombinant expression
CC   construct comprising at least one heterologous nucleic acid fragment
CC   operably linked to any one of the isolated nucleic acid fragments above,
CC   a plant comprising the recombinant expression construct in its genome and
CC   a method of regulating expression of at least one heterologous nucleotide
CC   sequence in the plant. The plants are dicotyledonous plants, preferably
CC   soybean. The method of regulating expression of at least one heterologous
CC   nucleotide sequence in a plant comprises transforming a plant cell with
CC   the recombinant expression construct, growing fertile mature plants from
CC   a transformed plant cell and selecting plants comprising a transformed
CC   plant cell expressing the heterologous nucleotide sequence.
CC   Alternatively, plants comprising a transformed plant cell expressing the
CC   heterologous nucleotide sequence during early seed development are
CC   selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC   and P34 promoters. The promoters are useful in regulating or promoting
CC   expression of at least one heterologous nucleic acid fragment in plants.
CC   This sequence represents a soybean annexin promoter truncated form of the
CC   invention.
XX
SQ   Sequence 1553 BP; 552 A; 237 C; 149 G; 615 T; 0 U; 0 Other;

  Query Match             77.2%;  Score 1553;  DB 13;  Length 1553;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-172;
  Matches 1553;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        460 CCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAA 519
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 CCCCGTACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAA 60

Qy        520 TATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTAT 579
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATTGTCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTAT 120

Qy        580 TAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATAT 639
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATAT 180

Qy        640 ACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAG 699
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 ACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAG 240

Qy        700 CTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATA 759
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATA 300

Qy        760 AGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTT 819
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTT 360

Qy        820 TTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGT 879
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGT 420

Qy        880 GAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAAACCTAAACAAACATGT 939
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAAACCTAAACAAACATGT 480

Qy        940 ATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTT 999
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 ATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTT 540

Qy       1000 TTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCT 1059
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCT 600

Qy       1060 CTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTAT 1119
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 CTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTAT 660

Qy       1120 CTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTC 1179
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTC 720

Qy       1180 AAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAA 1239
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 AAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAA 780

Qy       1240 AAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTT 1299
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 AAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTT 840

Qy       1300 TTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATC 1359
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATC 900

Qy       1360 TAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATA 1419
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATA 960

Qy       1420 TTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAAT 1479
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAAT 1020

Qy       1480 TTTAAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGC 1539
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 TTTAAATATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGC 1080
```

```
Qy       1540 ATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGA 1599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGA 1140

Qy       1600 TAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGA 1659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGA 1200

Qy       1660 AAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAA 1719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 AAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAA 1260

Qy       1720 GTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATA 1779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATA 1320

Qy       1780 GTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTG 1839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 GTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTG 1380

Qy       1840 CACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCC 1899
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1381 CACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCC 1440

Qy       1900 TGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAAT 1959
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1441 TGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAAT 1500

Qy       1960 GAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1501 GAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 1553


RESULT 6
ADR16030
ID   ADR16030 standard; DNA; 1367 BP.
XX
AC   ADR16030;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Soybean annexin promoter truncated sequence #4.
XX
KW   Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS   Glycine max.
XX
PN   US2004158052-A1.
XX
PD   12-AUG-2004.
XX
PF   11-FEB-2004; 2004US-00776889.
XX
PR   12-FEB-2003; 2003US-0446833P.
XX
PA   (KINN/) KINNEY A J.
PA   (LIUZ/) LIU Z.
XX
PI   Kinney AJ,  Liu Z;
XX
DR   WPI; 2004-592806/57.
XX
PT   New isolated nucleic acids comprising seed-specific soybean annexin and
PT   P34 soybean promoters, useful for regulating expression of at least one
PT   heterologous nucleic acid fragment in plants.
XX
PS   Claim 2; SEQ ID NO 16; 27pp; English.
XX
CC   The invention relates to an isolated nucleic acid fragment comprising a
CC   seed-specific soybean annexin promoter or seed-specific soybean P34
CC   promoter. The invention also relates to a recombinant expression
CC   construct comprising at least one heterologous nucleic acid fragment
CC   operably linked to any one of the isolated nucleic acid fragments above,
CC   a plant comprising the recombinant expression construct in its genome and
CC   a method of regulating expression of at least one heterologous nucleotide
CC   sequence in the plant. The plants are dicotyledonous plants, preferably
CC   soybean. The method of regulating expression of at least one heterologous
CC   nucleotide sequence in a plant comprises transforming a plant cell with
CC   the recombinant expression construct, growing fertile mature plants from
CC   a transformed plant cell and selecting plants comprising a transformed
CC   plant cell expressing the heterologous nucleotide sequence.
CC   Alternatively, plants comprising a transformed plant cell expressing the
CC   heterologous nucleotide sequence during early seed development are
CC   selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC   and P34 promoters. The promoters are useful in regulating or promoting
CC   expression of at least one heterologous nucleic acid fragment in plants.
CC   This sequence represents a soybean annexin promoter truncated form of the
CC   invention.
XX
SQ   Sequence 1367 BP; 493 A; 209 C; 137 G; 528 T; 0 U; 0 Other;

  Query Match           67.9%;  Score 1367;  DB 13;  Length 1367;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-150;
  Matches 1367;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        646 ATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATAT 705
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATAT 60

Qy        706 CATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAAT 765
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAAT 120

Qy        766 AAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCA 825
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCA 180

Qy        826 CATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAA 885
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAA 240

Qy        886 TAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTAC 945
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TAAGTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTAC 300

Qy        946 TCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTAT 1005
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTAT 360

Qy       1006 CATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTA 1065
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTA 420

Qy       1066 TTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTC 1125
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTC 480

Qy       1126 CACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGA 1185
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 CACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGA 540

Qy       1186 TTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCT 1245
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCT 600

Qy       1246 TTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTT 1305
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTT 660

Qy       1306 CTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATA 1365
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATA 720

Qy       1366 TATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTT 1425
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTT 780
```

```
Qy       1426 TATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAA 1485
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAA 840

Qy       1486 TATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTT 1545
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TATTTTTATATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTT 900

Qy       1546 TACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCG 1605
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCG 960

Qy       1606 TTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGA 1665
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 TTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGA 1020

Qy       1666 CGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCA 1725
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 CGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCA 1080

Qy       1726 ACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTA 1785
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 ACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTA 1140

Qy       1786 TTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATT 1845
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1141 TTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATT 1200

Qy       1846 TGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAA 1905
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1201 TGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAA 1260

Qy       1906 GGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTC 1965
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1261 GGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTC 1320

Qy       1966 ACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db       1321 ACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 1367
```

RESULT 7
ADR16031
ID ADR16031 standard; DNA; 1160 BP.
XX
AC ADR16031;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean annexin promoter truncated sequence #5.
XX
KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS Glycine max.
XX
PN US2004158052-A1.
XX
PD 12-AUG-2004.
XX
PF 11-FEB-2004; 2004US-00776889.
XX
PR 12-FEB-2003; 2003US-0446833P.
XX
PA (KINN/) KINNEY A J.
PA (LIUZ/) LIU Z.
XX
PI Kinney AJ, Liu Z;
XX
DR WPI; 2004-592806/57.
XX
PT New isolated nucleic acids comprising seed-specific soybean annexin and
PT P34 soybean promoters, useful for regulating expression of at least one
PT heterologous nucleic acid fragment in plants.
XX
PS Claim 2; SEQ ID NO 17; 27pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
CC seed-specific soybean annexin promoter or seed-specific soybean P34
CC promoter. The invention also relates to a recombinant expression
CC construct comprising at least one heterologous nucleic acid fragment
CC operably linked to any one of the isolated nucleic acid fragments above,
CC a plant comprising the recombinant expression construct in its genome and
CC a method of regulating expression of at least one heterologous nucleotide
CC sequence in the plant. The plants are dicotyledonous plants, preferably
CC soybean. The method of regulating expression of at least one heterologous
CC nucleotide sequence in a plant comprises transforming a plant cell with
CC the recombinant expression construct, growing fertile mature plants from
CC a transformed plant cell and selecting plants comprising a transformed
CC plant cell expressing the heterologous nucleotide sequence.
CC Alternatively, plants comprising a transformed plant cell expressing the
CC heterologous nucleotide sequence during early seed development are
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC and P34 promoters. The promoters are useful in regulating or promoting
CC expression of at least one heterologous nucleic acid fragment in plants.
CC This sequence represents a soybean annexin promoter truncated form of the
CC invention.
XX
SQ Sequence 1160 BP; 419 A; 190 C; 129 G; 422 T; 0 U; 0 Other;

Query Match 57.7%; Score 1160; DB 13; Length 1160;
Best Local Similarity 100.0%; Pred. No. 2.5e-126;
Matches 1160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy        853 ATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATT 912
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATT 60

Qy        913 TCCTAAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCA 972
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TCCTAAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCA 120

Qy        973 TTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTA 1032
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTA 180

Qy       1033 CTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAAT 1092
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAAT 240

Qy       1093 TGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAA 1152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAA 300

Qy       1153 AGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATT 1212
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 AGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATT 360

Qy       1213 GATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAAT 1272
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAAT 420

Qy       1273 CTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTT 1332
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTT 480

Qy       1333 TTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTA 1392
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTA 540

Qy       1393 AACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATA 1452
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATA 600

Qy       1453 AGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAA 1512
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        601 AGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAA 660

Qy       1513 TATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCT 1572
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCT 720

Qy       1573 ATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTC 1632
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 ATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTC 780

Qy       1633 TTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACT 1692
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACT 840

Qy       1693 CTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTT 1752
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 CTTCACACCATTTTTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTT 900

Qy       1753 TCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTTCACATGATTTATTAACGT 1812
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 TCTTTTTACAACACTTTCTTTGAAATAGTAGTATTTTTTTTTTCACATGATTTATTAACGT 960

Qy       1813 GCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGA 1872
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        961 GCCAAAAGATGCTTATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGA 1020

Qy       1873 AAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCA 1932
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1021 AAAAGCATTGTTCTAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCA 1080

Qy       1933 CTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTT 1992
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1081 CTATATAAATTGTGACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTT 1140

Qy       1993 AACTAAGAAGCAATACTTCA 2012
              ||||||||||||||||||||
Db       1141 AACTAAGAAGCAATACTTCA 1160


RESULT 8
ADR16032
ID   ADR16032 standard; DNA; 967 BP.
XX
AC   ADR16032;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Soybean annexin promoter truncated sequence #6.
XX
KW   Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS   Glycine max.
XX
PN   US2004158052-A1.
XX
PD   12-AUG-2004.
XX
PF   11-FEB-2004; 2004US-00776889.
XX
PR   12-FEB-2003; 2003US-0446833P.
XX
PA   (KINN/) KINNEY A J.
PA   (LIUZ/) LIU Z.
XX
PI   Kinney AJ,  Liu Z;
XX
DR   WPI; 2004-592806/57.
XX
PT   New isolated nucleic acids comprising seed-specific soybean annexin and
PT   P34 soybean promoters, useful for regulating expression of at least one
PT   heterologous nucleic acid fragment in plants.
XX
PS   Claim 2; SEQ ID NO 18; 27pp; English.
XX
CC   The invention relates to an isolated nucleic acid fragment comprising a
CC   seed-specific soybean annexin promoter or seed-specific soybean P34
CC   promoter. The invention also relates to a recombinant expression
CC   construct comprising at least one heterologous nucleic acid fragment
CC   operably linked to any one of the isolated nucleic acid fragments above,
CC   a plant comprising the recombinant expression construct in its genome and
CC   a method of regulating expression of at least one heterologous nucleotide
CC   sequence in the plant. The plants are dicotyledonous plants, preferably
CC   soybean. The method of regulating expression of at least one heterologous
CC   nucleotide sequence in a plant comprises transforming a plant cell with
CC   the recombinant expression construct, growing fertile mature plants from
CC   a transformed plant cell and selecting plants comprising a transformed
CC   plant cell expressing the heterologous nucleotide sequence.
CC   Alternatively, plants comprising a transformed plant cell expressing the
CC   heterologous nucleotide sequence during early seed development are
CC   selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC   and P34 promoters. The promoters are useful in regulating or promoting
CC   expression of at least one heterologous nucleic acid fragment in plants.
CC   This sequence represents a soybean annexin promoter truncated form of the
CC   invention.
XX
SQ   Sequence 967 BP; 357 A; 161 C; 112 G; 337 T; 0 U; 0 Other;

  Query Match             48.1%;  Score 967;  DB 13;  Length 967;
  Best Local Similarity   100.0%;  Pred. No. 6.2e-104;
  Matches  967;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1046 TTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTT 1105
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTT 60

Qy       1106 TATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGT 1165
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGT 120

Qy       1166 CTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATC 1225
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATC 180

Qy       1226 TATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAA 1285
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAA 240

Qy       1286 AGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTA 1345
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTA 300

Qy       1346 TATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAA 1405
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAA 360

Qy       1406 TATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATA 1465
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATA 420

Qy       1466 TGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATA 1525
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATA 480

Qy       1526 AATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTG 1585
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTG 540

Qy       1586 AAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTT 1645
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 AAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTT 600

Qy       1646 TTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTT 1705
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 TTCATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTT 660

Qy       1706 TTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACA 1765
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 TTACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACA 720

Qy       1766 CTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCT 1825
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 CTTTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCT 780

Qy       1826 TATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTC 1885
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        781 TATTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTC 840

Qy       1886 TAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGT 1945
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        841 TAACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGT 900

Qy       1946 GACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAA 2005
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        901 GACGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAA 960

Qy       2006 TACTTCA 2012
                |||||||
Db        961 TACTTCA 967
```

RESULT 9
ADR16033
ID ADR16033 standard; DNA; 770 BP.
XX
AC ADR16033;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean annexin promoter truncated sequence #7.
XX
KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS Glycine max.
XX
PN US2004158052-A1.
XX
PD 12-AUG-2004.
XX
PF 11-FEB-2004; 2004US-00776889.
XX
PR 12-FEB-2003; 2003US-0446833P.
XX
PA (KINN/) KINNEY A J.
PA (LIUZ/) LIU Z.
XX
PI Kinney AJ, Liu Z;
XX
DR WPI; 2004-592806/57.
XX
PT New isolated nucleic acids comprising seed-specific soybean annexin and
PT P34 soybean promoters, useful for regulating expression of at least one
PT heterologous nucleic acid fragment in plants.
XX
PS Claim 2; SEQ ID NO 19; 27pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
CC seed-specific soybean annexin promoter or seed-specific soybean P34
CC promoter. The invention also relates to a recombinant expression
CC construct comprising at least one heterologous nucleic acid fragment
CC operably linked to any one of the isolated nucleic acid fragments above,
CC a plant comprising the recombinant expression construct in its genome and
CC a method of regulating expression of at least one heterologous nucleotide
CC sequence in the plant. The plants are dicotyledonous plants, preferably
CC soybean. The method of regulating expression of at least one heterologous
CC nucleotide sequence in a plant comprises transforming a plant cell with
CC the recombinant expression construct, growing fertile mature plants from
CC a transformed plant cell and selecting plants comprising a transformed
CC plant cell expressing the heterologous nucleotide sequence.
CC Alternatively, plants comprising a transformed plant cell expressing the
CC heterologous nucleotide sequence during early seed development are
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC and P34 promoters. The promoters are useful in regulating or promoting
CC expression of at least one heterologous nucleic acid fragment in plants.
CC This sequence represents a soybean annexin promoter truncated form of the
CC invention.
XX
SQ Sequence 770 BP; 303 A; 108 C; 89 G; 270 T; 0 U; 0 Other;

Query Match 38.3%; Score 770; DB 13; Length 770;
Best Local Similarity 100.0%; Pred. No. 4.5e-81;
Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1243 TCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTT 1302
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTT 60

Qy       1303 TTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAA 1362
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAA 120

Qy       1363 ATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTT 1422
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTT 180

Qy       1423 TTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTT 1482
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTT 240

Qy       1483 AAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATA 1542
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAATATTTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATA 300

Qy       1543 TTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAA 1602
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TTTTACAACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAA 360

Qy       1603 TCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAG 1662
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAG 420

Qy       1663 GGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTC 1722
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTC 480

Qy       1723 TCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTA 1782
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 TCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTA 540

Qy       1783 GTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCAC 1842
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCAC 600

Qy       1843 ATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGT 1902
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 ATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGT 660

Qy       1903 GAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAA 1962
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAA 720

Qy       1963 TTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
                |||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 TTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 770
```

RESULT 10
ADR16034
ID ADR16034 standard; DNA; 584 BP.
XX
AC ADR16034;
XX
DT 18-NOV-2004 (first entry)

XX
DE Soybean annexin promoter truncated sequence #8.
XX
KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS Glycine max.
XX
PN US2004158052-A1.
XX
PD 12-AUG-2004.
XX
PF 11-FEB-2004; 2004US-00776889.
XX
PR 12-FEB-2003; 2003US-0446833P.
XX
PA (KINN/) KINNEY A J.
PA (LIUZ/) LIU Z.
XX
PI Kinney AJ, Liu Z;
XX
DR WPI; 2004-592806/57.
XX
PT New isolated nucleic acids comprising seed-specific soybean annexin and
PT P34 soybean promoters, useful for regulating expression of at least one
PT heterologous nucleic acid fragment in plants.
XX
PS Claim 2; SEQ ID NO 20; 27pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
CC seed-specific soybean annexin promoter or seed-specific soybean P34
CC promoter. The invention also relates to a recombinant expression
CC construct comprising at least one heterologous nucleic acid fragment
CC operably linked to any one of the isolated nucleic acid fragments above,
CC a plant comprising the recombinant expression construct in its genome and
CC a method of regulating expression of at least one heterologous nucleotide
CC sequence in the plant. The plants are dicotyledonous plants, preferably
CC soybean. The method of regulating expression of at least one heterologous
CC nucleotide sequence in a plant comprises transforming a plant cell with
CC the recombinant expression construct, growing fertile mature plants from
CC a transformed plant cell and selecting plants comprising a transformed
CC plant cell expressing the heterologous nucleotide sequence.
CC Alternatively, plants comprising a transformed plant cell expressing the
CC heterologous nucleotide sequence during early seed development are
CC selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC and P34 promoters. The promoters are useful in regulating or promoting
CC expression of at least one heterologous nucleic acid fragment in plants.
CC This sequence represents a soybean annexin promoter truncated form of the
CC invention.
XX
SQ Sequence 584 BP; 227 A; 95 C; 78 G; 184 T; 0 U; 0 Other;

Query Match 29.0%; Score 584; DB 13; Length 584;
Best Local Similarity 100.0%; Pred. No. 1.8e-59;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1429 ATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATAT 1488
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATAT 60

Qy       1489 TTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTAC 1548
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTTTATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTAC 120

Qy       1549 AACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTA 1608
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 AACCTTAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTA 180

Qy       1609 AGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGA 1668
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGAAAGGGACGA 240

Qy       1669 AAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACA 1728
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AAAACCACCTCACCATGAATCACTCTTCACACCATTTTTACTAGCAAACAAGTCTCAACA 300

Qy       1729 ACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTT 1788
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACTTTCTTTGAAATAGTAGTATTT 360

Qy       1789 TTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGT 1848
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTATTGAATAGAGTGCACATTTGT 420

Qy       1849 AATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGC 1908
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATCCTGTGAAGGC 480

Qy       1909 GTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACA 1968
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 GTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAATGAATTCACA 540

Qy       1969 TAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             |||||||||||||||||||||||||||||||||||||||||||||
Db        541 TAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 584
```

RESULT 11
ADR16035
ID ADR16035 standard; DNA; 425 BP.
XX
AC ADR16035;
XX
DT 18-NOV-2004 (first entry)
XX
DE Soybean annexin promoter truncated sequence #9.
XX
KW Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS Glycine max.
XX
PN US2004158052-A1.
XX
PD 12-AUG-2004.
XX
PF 11-FEB-2004; 2004US-00776889.
XX
PR 12-FEB-2003; 2003US-0446833P.
XX
PA (KINN/) KINNEY A J.
PA (LIUZ/) LIU Z.
XX
PI Kinney AJ, Liu Z;
XX
DR WPI; 2004-592806/57.
XX
PT New isolated nucleic acids comprising seed-specific soybean annexin and
PT P34 soybean promoters, useful for regulating expression of at least one
PT heterologous nucleic acid fragment in plants.
XX
PS Claim 2; SEQ ID NO 21; 27pp; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
CC seed-specific soybean annexin promoter or seed-specific soybean P34
CC promoter. The invention also relates to a recombinant expression
CC construct comprising at least one heterologous nucleic acid fragment
CC operably linked to any one of the isolated nucleic acid fragments above,
CC a plant comprising the recombinant expression construct in its genome and
CC a method of regulating expression of at least one heterologous nucleotide
CC sequence in the plant. The plants are dicotyledonous plants, preferably
CC soybean. The method of regulating expression of at least one heterologous
CC nucleotide sequence in a plant comprises transforming a plant cell with
CC the recombinant expression construct, growing fertile mature plants from
CC a transformed plant cell and selecting plants comprising a transformed
CC plant cell expressing the heterologous nucleotide sequence.
CC Alternatively, plants comprising a transformed plant cell expressing the
CC heterologous nucleotide sequence during early seed development are

```
CC   selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC   and P34 promoters. The promoters are useful in regulating or promoting
CC   expression of at least one heterologous nucleic acid fragment in plants.
CC   This sequence represents a soybean annexin promoter truncated form of the
CC   invention.
XX
SQ   Sequence 425 BP; 154 A; 81 C; 64 G; 126 T; 0 U; 0 Other;

  Query Match            21.1%;  Score 425;  DB 13;  Length 425;
  Best Local Similarity  100.0%;  Pred. No. 5.1e-41;
  Matches  425;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1588 ATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTT 1647
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTCTTGCCATTTGTTTTT 60

Qy       1648 CATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTT 1707
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CATGCAAACAGAAAGGGACGAAAAACCACCTCACCATGAATCACTCTTCACACCATTTTT 120

Qy       1708 ACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACT 1767
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACTAGCAAACAAGTCTCAACAACTGAAGCCAGCTCTCTTTCCGTTTCTTTTTACAACACT 180

Qy       1768 TTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTA 1827
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 TTCTTTGAAATAGTAGTATTTTTTTTTCACATGATTTATTAACGTGCCAAAAGATGCTTA 240

Qy       1828 TTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTA 1887
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TTGAATAGAGTGCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTA 300

Qy       1888 ACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGA 1947
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 ACACGATAATCCTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGA 360

Qy       1948 CGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATA 2007
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CGAAAGCAAAATGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATA 420

Qy       2008 CTTCA 2012
             |||||
Db        421 CTTCA 425


RESULT 12
ADR16036
ID   ADR16036 standard; DNA; 174 BP.
XX
AC   ADR16036;
XX
DT   18-NOV-2004  (first entry)
XX
DE   Soybean annexin promoter truncated sequence #10.
XX
KW   Soybean; annexin promoter; ds; dicotyledonous plant; plant.
XX
OS   Glycine max.
XX
PN   US2004158052-A1.
XX
PD   12-AUG-2004.
XX
PF   11-FEB-2004; 2004US-00776889.
XX
PR   12-FEB-2003; 2003US-0446833P.
XX
PA   (KINN/) KINNEY A J.
PA   (LIUZ/) LIU Z.
XX
PI   Kinney AJ,  Liu Z;
XX
DR   WPI; 2004-592806/57.
XX
PT   New isolated nucleic acids comprising seed-specific soybean annexin and
PT   P34 soybean promoters, useful for regulating expression of at least one
PT   heterologous nucleic acid fragment in plants.
XX
PS   Claim 2; SEQ ID NO 22; 27pp; English.
XX
CC   The invention relates to an isolated nucleic acid fragment comprising a
CC   seed-specific soybean annexin promoter or seed-specific soybean P34
CC   promoter. The invention also relates to a recombinant expression
CC   construct comprising at least one heterologous nucleic acid fragment
CC   operably linked to any one of the isolated nucleic acid fragments above,
CC   a plant comprising the recombinant expression construct in its genome and
CC   a method of regulating expression of at least one heterologous nucleotide
CC   sequence in the plant. The plants are dicotyledonous plants, preferably
CC   soybean. The method of regulating expression of at least one heterologous
CC   nucleotide sequence in a plant comprises transforming a plant cell with
CC   the recombinant expression construct, growing fertile mature plants from
CC   a transformed plant cell and selecting plants comprising a transformed
CC   plant cell expressing the heterologous nucleotide sequence.
CC   Alternatively, plants comprising a transformed plant cell expressing the
CC   heterologous nucleotide sequence during early seed development are
CC   selected. The nucleic acid is useful as a plant promoter, e.g. annexin
CC   and P34 promoters. The promoters are useful in regulating or promoting
CC   expression of at least one heterologous nucleic acid fragment in plants.
CC   This sequence represents a soybean annexin promoter truncated form of the
CC   invention.
XX
SQ   Sequence 174 BP; 72 A; 28 C; 30 G; 44 T; 0 U; 0 Other;

  Query Match             8.6%;  Score 174;  DB 13;  Length 174;
  Best Local Similarity   100.0%;  Pred. No. 7.6e-12;
  Matches  174;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1839 GCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATC 1898
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GCACATTTGTAATGTACTACTAATTAGAACATGAAAAAGCATTGTTCTAACACGATAATC 60

Qy       1899 CTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAA 1958
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGTGAAGGCGTTAACTCCAAAGATCCAATTTCACTATATAAATTGTGACGAAAGCAAAA 120

Qy       1959 TGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 2012
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 TGAATTCACATAGCTGAGAGAGAAAGGAAAGGTTAACTAAGAAGCAATACTTCA 174


RESULT 13
ABL34155
ID   ABL34155 standard; DNA; 15548 BP.
XX
AC   ABL34155;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human immune system associated gene SEQ ID NO: 2128.
XX
KW   Human; immune system disease; cytosine methylation; antiasthmatic;
KW   antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW   neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW   antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW   antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW   acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW   neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
KW   ds.
XX
OS   Homo sapiens.
XX
PN   WO200200928-A2.
XX
PD   03-JAN-2002.
XX
PF   02-JUL-2001; 2001WO-EP007537.
```

```
XX
PR   30-JUN-2000; 2000DE-01032529.
PR   01-SEP-2000; 2000DE-01043826.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Olek A,  Piepenbrock C,  Berlin K;
XX
DR   WPI; 2002-130909/17.
XX
PT   Nucleic acid comprising fragment of chemically modified gene, useful for
PT   diagnosis and treatment of diseases associated with abnormal cytosine
PT   methylation.
XX
PS   Claim 1; SEQ ID NO 2128; 32pp + Sequence Listing; German.
XX
CC   The present invention provides a number of human immune system associated
CC   genes which are modified by the methylation of cytosines. The sequences
CC   can be used in the diagnosis and treatment of immune system disorders,
CC   including eye diseases such as retinopathy, neovascular glaucoma and
CC   macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC   leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC   rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC   diseases. The present sequence is a gene of the invention
XX
SQ   Sequence 15548 BP; 4209 A; 247 C; 2903 G; 8189 T; 0 U; 0 Other;

  Query Match             7.8%;  Score 156;  DB 6;  Length 15548;
  Best Local Similarity   50.4%;  Pred. No. 4.5e-10;
  Matches  594;  Conservative    0;  Mismatches  565;  Indels   19;  Gaps    8;

Qy        495 TATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTAT 554
              ||| ||||| ||  || || |   |  | |    | |  || ||    |  |  ||| |
Db      11586 TATATTTTATTTTTATATTTTATTTTTATATTTTATTTTTATATTTTATTTTTATATTTT 11645

Qy        555 TATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGT 614
                | |||  ||||    |  | |  |  || ||||||  | |||||| | |  |||| |
Db      11646 ATTTTTATATTTTATATTTATTTTATTTTTATATTTTATATTTATTTTATTTTTATATTT 11705

Qy        615 AAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGA 674
              |  |  ||| |||||||    | || ||  ||||||    |||||    |       |
Db      11706 TATATTTATTTTATTTTTATATTTTATTTTTATATTTTATTATTATATTTTATTTTTATA 11765

Qy        675 TATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTT 734
              | | |    | |||| | |||  |  | |    |||    | || ||||   |||  |||
Db      11766 TTTTATTTTTATATTTTATTTTTATATTTTATTTTTATATTTTATTTTTATATTTTATTT 11825

Qy        735 CTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTT 794
              | |  |||  | ||  |||| ||    | |  |||||   | ||| |||||||   |||
Db      11826 TTATATTTTATTTTTATATTTTATTTTTATTTTATTTTTATATTTTATATTTAT---TTT 11882

Qy        795 ATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTAT 854
              ||    ||||||||  ||   | | |||| | |||||  ||  | |  | || |   |||
Db      11883 ATTTTTATATTTTATATTTATTTTATTTTTATATTTTATTTTTATATTTATTTTTATAT 11942

Qy        855 GATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTC 914
              | |       | | |  ||  | | |   |  | |    | | | | | |   || | |
Db      11943 GTTATTTTTATATGTTATATATATTTTATTTTTATATGTTATATATATTTTATTTTTATA 12002

Qy        915 CTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATT 974
              | | |   |  |    || | |    | | ||| | || | | | |  ||   ||
Db      12003 TGTTATATATATTTTATTTTTATATTTTATTTTTATATGTTATTTTTATAT-GTTATTTT 12061

Qy        975 TTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACT 1034
              | | | ||  | |  |   |||||||| ||| |||  ||  |  ||| || | | |  |
Db      12062 TATATGTTATTTTTATATGTTATTTTTATATGTTATTTTTATATGTTATTTTTATATGTT 12121

Qy       1035 CTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTG 1094
              | ||||| | || | | |||   |   || | |       |||||||   |  | | |
Db      12122 ATTTTTATATGTTATATATATATGTTATTTTTATAGTTATATATATATGTTATATATATA 12181

Qy       1095 GGGCATGCCTTTATCACTACTCTATCTCCTCCA---CTAAATCATTTAAATGAAACTGAA 1151
                |  ||   | |||      | ||| |  |  |    ||  | || || ||  |  | |
```

```
Db      12182 TGTTATATATATATGTTATATATATATGTTATATATATATGTTATATATATATATGTTAT 12241

Qy       1152 AAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGAT 1211
              |   ||  |  |       |       |  |   |   |  |    ||  |      ||
Db      12242 ATATATATGTTATATATATATATGTTATATATATATGTTATATATATATATGTTATATAT 12301

Qy       1212 TGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAA 1271
                ||     ||||| |||    | | || |  | ||       ||| ||   |||  ||  |
Db      12302 ATATGTTATATATATATATGTTATATATATATGTTATATATATATAT-GTTATATATATA 12360

Qy       1272 TCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGT-TATT 1330
              | ||| ||| ||| | | || |    | | | || |  ||| |   | | |||| | ||||
Db      12361 TGTTATATATATATATGTTATATATATATATGTTATGTATATATGTTATATATATATATT 12420

Qy       1331 TTTTGTTTAGATGTATATTCGAATAAATCTAAATATATG--------ATAATGATTTTTT 1382
              || ||| || ||||    |  |   | | ||  |||||         |||   ||||| |
Db      12421 TTATGTATATATGTTATATATATATATTTTATGTATATATTTTATATATATATATTTTAT 12480

Qy       1383 ATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACA 1442
                |||  |||  | |||| ||  || |||  |||    ||| | ||||||  || ||   |
Db      12481 GTATATATTTTATATAT-ATATATTTTATGTATATATTTTATATATATATATTTTATGTA 12539

Qy       1443 TA-ATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACG 1501
              || |||||||   ||| |   ||||   | | |  || |  |     ||||| | ||
Db      12540 TATATTTTATATATATATATTTTATGTATATATTTTATATATATATATTTTATGTATATA 12599

Qy       1502 AGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATT 1561
                | | | | |||| |||   ||| ||     |    |||||||| |    || ||  |
Db      12600 TGTTATATATATATATTATATATATATTTCGTATATATATATTTTATATATATATTATAT 12659

Qy       1562 CATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCC 1621
                | | ||| ||||| ||||||| | ||  | |  | |||    ||  |   |   ||
Db      12660 ATTTATTTCGTATATATATATTTTATATATATATTATATATATATTTCGTATATATATAT 12719

Qy       1622 TACGTCATCTCTTGCCATTTGTTTTTCATGCAAACAGA 1659
              |   |       |   || | | |||| |  | | | |
Db      12720 TTTATATATATATTATATATATATTTCGTATATATATA 12757


RESULT 14
ABZ10246/c
ID   ABZ10246 standard; DNA; 8056 BP.
XX
AC   ABZ10246;
XX
DT   16-JAN-2003  (first entry)
XX
DE   Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
KW   Human; haematopoietic cell proliferation disorder; cytostatic;
KW   gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW   cytosine methylation state; gene; ds.
XX
OS   Homo sapiens.
XX
PN   WO200277272-A2.
XX
PD   03-OCT-2002.
XX
PF   26-MAR-2002; 2002WO-EP003401.
XX
PR   26-MAR-2001; 2001US-0278333P.
XX
PA   (EPIG-) EPIGENOMICS AG.
XX
PI   Berlin K,  Braun A,  Distler J,  Guetig D,  Howe A,  Mueller J;
PI   Olek A,  Piepenbrock C,  Adorjan P,  Grabs G,  Lesche R,  Leu E;
PI   Lewin A,  Lipscher E,  Maier S,  Model F,  Mueller V,  Otto T,  Pelet C;
PI   Schwope I,  Ziebarth H;
XX
DR   WPI; 2003-018942/01.
XX
PT   Detecting and differentiating between hematopoietic cell proliferative
```

PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 386; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

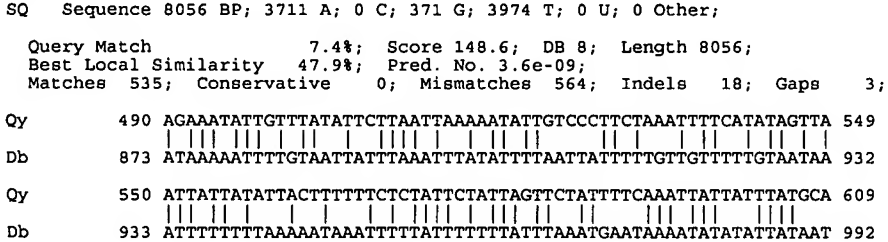
Query Match 7.5%; Score 150.2; DB 8; Length 8056;
Best Local Similarity 45.5%; Pred. No. 2.3e-09;
Matches 725; Conservative 0; Mismatches 858; Indels 10; Gaps 5;

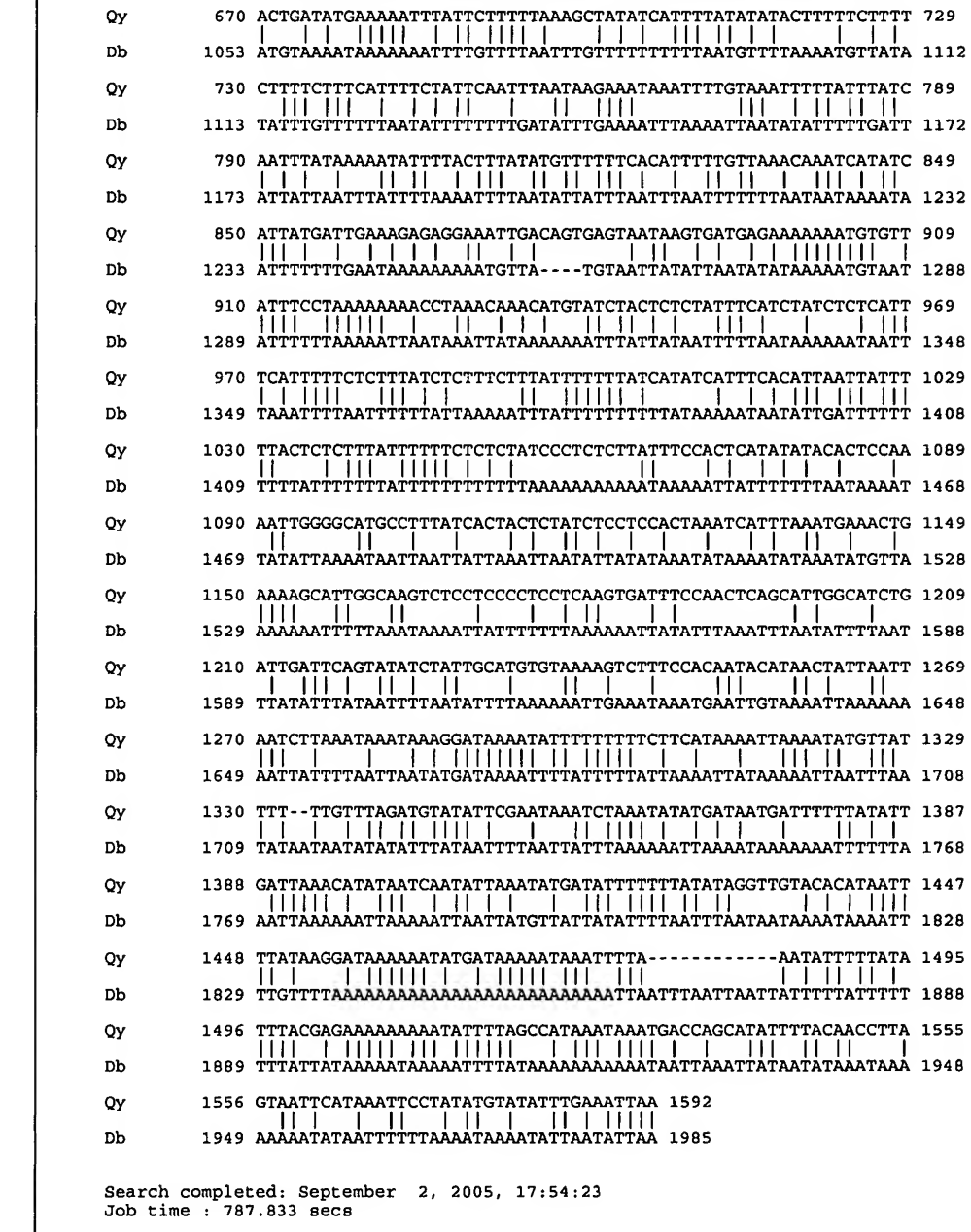
```
Qy        16 TTATATGGTGTTTAGATGGATTCACATGCAAGTTTTTATTTCAATCCCTTTTCCTTTGAA 75
Db      2543 TTATTTTATAATTTAATTTATTTTTTTTTATTTTTCATTAATATTTTAATTATTTTAAAA 2484

Qy        76 TAACTGACCAAGAACAACAAGAAAAAAAAAAAAAAGAAAAGGATCATTTTGAAAGGATATT 135
Db      2483 CAAAATAATAAAAAATATATAAATAAAAATAAAAAAAAATTAATTTTTAAAAATAAATAA 2424

Qy       136 TTTCGCTCCTATTCAAATACTGTATTTTTACCAAAAAAACTGTATT-TTTCCTACACTCT 194
Db      2423 ATTATTTTTTTTTTAATTTTCAAAAAATAATAAAATTTAATAAATTATTTATAAAAATAA 2364

Qy       195 CAAGCTTTGTTTTTCGCTTCGACTCTCATGATTTCCTTCATATGCCAATCACTCTATTTA 254
Db      2363 AAAATTATATTTAAAAATTAAATAAAATTTATTTAAATACAAAAATTAATAAATTATTTT 2304

Qy       255 TAAATGGCATAAGGTAGTGTGAACAATTGCAAAGCTTGTCATCAAAAGCTTGCAATGTAC 314
Db      2303 AAAAATAAATAAATTAAAAAATTAAAATTTTTTTAATTTATTTAAATATTAAAAAATAA 2244

Qy       315 AAATTAATGTTTTTCATGCCTTTCAAAATTATCTGCACCCCCTAGCTATTAATCTAACAT 374
Db      2243 ATTTTTATTAAAATTAAATTTTTTTTTTTTTTTATTTTTTTTTTCATTTTTTTAAATTTT 2184

Qy       375 CTAAGTAAGGCTAGTGAATTTTTTCGAATAGTCATGCAGTGCATTAATTTCCCCGTGACT 434
Db      2183 TAATAAAATTATAATTTTTTTTTTTTTTATTTTTATTTTTAATAAAATAAAAAATTATT 2124

Qy       435 ATTTTGGCTTTGACTCCAACACTGGCCCCGTACATCCGTCCCTCATTACATGAAAAGAAA 494
Db      2123 AAAATTATATTAAATAATTTATAATTATTATTTATAATATTTAAAATATTTTTAATAAAA 2064

Qy       495 TATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTAT 554
Db      2063 TTTTTTAAAAATAAAAAATATTTAAAAATATATTTT---ATTTTAAATATATTTAAAATT 2007

Qy       555 TATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGT 614
Db      2006 TTATTAAATTTTAAATAAAAATTAATATTAATATTTTATTTTAAAAAATTATATTTTTTT 1947

Qy       615 AAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGA 674
Db      1946 TATTTATATTATAATTTAATTATTTTTTTTTTATAAAATTTTTATTTTTATAATAAAAA 1887

Qy       675 TATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTT 734
Db      1886 AAATAAAAATAATTAATTAAATTAATTTTTTTTTTTTTTTTTTTTTTTTTAAAACAAAA 1827

Qy       735 CTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTT 794
Db      1826 TTTTATTTTATTATTAAATTAAAATATAATAACATAATTAATTTTTAATTTTTTAATTTA 1767

Qy       795 ATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTAT 854
Db      1766 AAAAATTTTTTTATTTTAATTTTTTAAATAATTAAAATTATAAATATATATTATTATATT 1707

Qy       855 GATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTC 914
Db      1706 AAATTAATTTTTATAATTTTAATAAAAATAAAATTTTATCATATTAATTAAAATAATTTT 1647

Qy       915 CTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATT 974
Db      1646 TTTTAATTTTACAATTCATTTATTTCAATTTTTTAAAATATTAAAATTATAAATATAAAT 1587

Qy       975 TTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATT--AATTATTTTTA 1032
Db      1586 TAAAATATTAAATTTAAATATAATTTTTTAAAAAAATAATTTTATTTAAAAATTTTTTTA 1527

Qy      1033 CTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAAT 1092
Db      1526 ACATATTTATATTTTATATTTATATAATATTAATTTAATAATTAATTATTTTAATATAAT 1467

Qy      1093 TGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAA 1152
Db      1466 TTTATTAAAAAAATAATTTTTATTTTTTTTTTTAAAAAAAAAAAATAAAAAAATAAAAAA 1407

Qy      1153 AGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATT 1212
Db      1406 AAAATCAATATTATTTTTATAAAAAAAAAATAAATTTTTAATAAAAAATTAAAATTTA-- 1349

Qy      1213 GATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAAT 1272
Db      1348 AATTATTTTTTATTAAAAATTATAATAAATTTTTTTATAATTTATTAATTTTTAAAAAAT 1289

Qy      1273 CTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTT 1332
Db      1288 ATTACATTTTTATATATTAATATAATTACATAACATTTTTTTTTTATTCAAAAAAATTATT 1229

Qy      1333 TTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTA 1392
Db      1228 TTATTATTAAAAAAATTAAATTAAATAATATTAAAATTTTAAAATAAATTAATAATAATC 1169

Qy      1393 AACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATA 1452
Db      1168 AAAAATATATTAATTTTAAATTTTCAAATATCAAAAAAAATATTAAAAAACAAATATATA 1109

Qy      1453 AGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAAAAA 1512
Db      1108 ACATTTTAAAACATTAAAAAAAAAACAAATTAAAACAAAATTTTTTTATTTTACATAATT 1049

Qy      1513 TATTTTAGCCATAAATAAATGACCAGCATATTTT--ACAACCTTAGTAATTCATAAATTC 1570
Db      1048 AAATTTCAAAATTAAAAAATAATAAAAAAATTTTTAAAATAATTATTTTTAAAAAAATTA 989

Qy      1571 CTATATGTATATTTGAAATTAAAAACAGATAAT 1603
Db       988 TAATATATATTTTATTCATTTAAATAAAAAAT 956
```

RESULT 15
ABZ10246
ID ABZ10246 standard; DNA; 8056 BP.

XX
AC ABZ10246;
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Distler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E;
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwope I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 386; 117pp; English.
XX
CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ11118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used: for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related DNA
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative
CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;

Query Match 7.4%; Score 148.6; DB 8; Length 8056;
Best Local Similarity 47.9%; Pred. No. 3.6e-09;
Matches 535; Conservative 0; Mismatches 564; Indels 18; Gaps 3;





Search completed: September 2, 2005, 17:54:23
Job time : 787.833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2005, 10:34:39 ; Search time 4923.25 Seconds
(without alignments)
15555.832 Million cell updates/sec

Title: US-10-776-889-1
Perfect score: 2012
Sequence: 1 atcttaggcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_htc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	172.4	8.6	1626	7	CF238805	CF238805 AGENCOURT
	2	165	8.2	1348	9	CG749499	CG749499 P043-4-A0
c	3	159.8	7.9	1608	9	CL118721	CL118721 ISB1-72J8
	4	159.2	7.9	1242	9	CL068807	CL068807 CH216-115
	5	156	7.8	1542	9	AG386981	AG386981 Mus muscu
	6	155.8	7.7	1896	9	CG753083	CG753083 P048-1-C0
	7	155.4	7.7	1378	9	AG350209	AG350209 Mus muscu
	8	154.4	7.7	1210	9	CG749728	CG749728 P044-1-C0
	9	153	7.6	1805	9	CL080711	CL080711 CH216-159
	10	152.8	7.6	1392	9	CG757503	CG757503 P052-4-C0
c	11	152.4	7.6	1632	9	CL082569	CL082569 CH216-167
	12	152	7.6	1594	9	CL110653	CL110653 ISB1-53P2
c	13	150.6	7.5	1594	9	CL038406	CL038406 CH216-46A
	14	149.6	7.4	1981	9	CL082000	CL082000 CH216-165
	15	149	7.4	1074	8	BZ696936	BZ696936 SP__Ba009
	16	149	7.4	1391	9	CG754863	CG754863 P050-2-G0
c	17	148.8	7.4	1241	9	AG448181	AG448181 Mus muscu
	18	148.4	7.4	1539	9	AG340947	AG340947 Mus muscu
c	19	148.4	7.4	1599	9	CL083840	CL083840 ISB1-2H14
	20	148	7.4	2157	9	CL081966	CL081966 CH216-165
	21	147.6	7.3	2087	9	AG333887	AG333887 Mus muscu
c	22	145.6	7.2	1745	9	AG338221	AG338221 Mus muscu
c	23	145.2	7.2	1021	9	CNS014DY	AL104032 Drosophil
	24	145.2	7.2	1227	9	AG430010	AG430010 Mus muscu
c	25	144.8	7.2	1260	9	CL491610	CL491610 SAIL_559_
c	26	144	7.2	1217	9	CL062848	CL062848 CH216-98N
	27	144	7.2	1354	9	CG744717	CG744717 P037-2-F0
	28	143.8	7.1	1489	9	AG350139	AG350139 Mus muscu
c	29	143.4	7.1	1277	8	CC253231	CC253231 CH261-180
	30	143.4	7.1	1313	7	CK997149	CK997149 ip15c02.b
	31	143.4	7.1	1454	9	CG747614	CG747614 P041-2-A0
	32	142.6	7.1	1268	9	AG347098	AG347098 Mus muscu
	33	142.4	7.1	1202	8	CC262481	CC262481 CH261-167
c	34	142.4	7.1	1224	9	CL077121	CL077121 CH216-143
	35	141.6	7.0	1503	9	AG346761	AG346761 Mus muscu
c	36	140.2	7.0	1205	9	CL143963	CL143963 ISB1-123O
	37	140.2	7.0	1928	9	CL073845	CL073845 CH216-130
c	38	140	7.0	1536	9	CL078538	CL078538 CH216-151
	39	139	6.9	1355	9	AG346348	AG346348 Mus muscu
c	40	138.4	6.9	1493	9	CL078589	CL078589 CH216-151
	41	137.6	6.8	1507	9	AG346189	AG346189 Mus muscu
	42	137.2	6.8	1029	9	CNS01ZGM	AL174271 Tetraodon
	43	137.2	6.8	1453	9	AJ591978	AJ591978 Arabidops
c	44	137	6.8	888	8	AZ549422	AZ549422 ENTDP49TF
	45	137	6.8	1211	9	AG349657	AG349657 Mus muscu

ALIGNMENTS

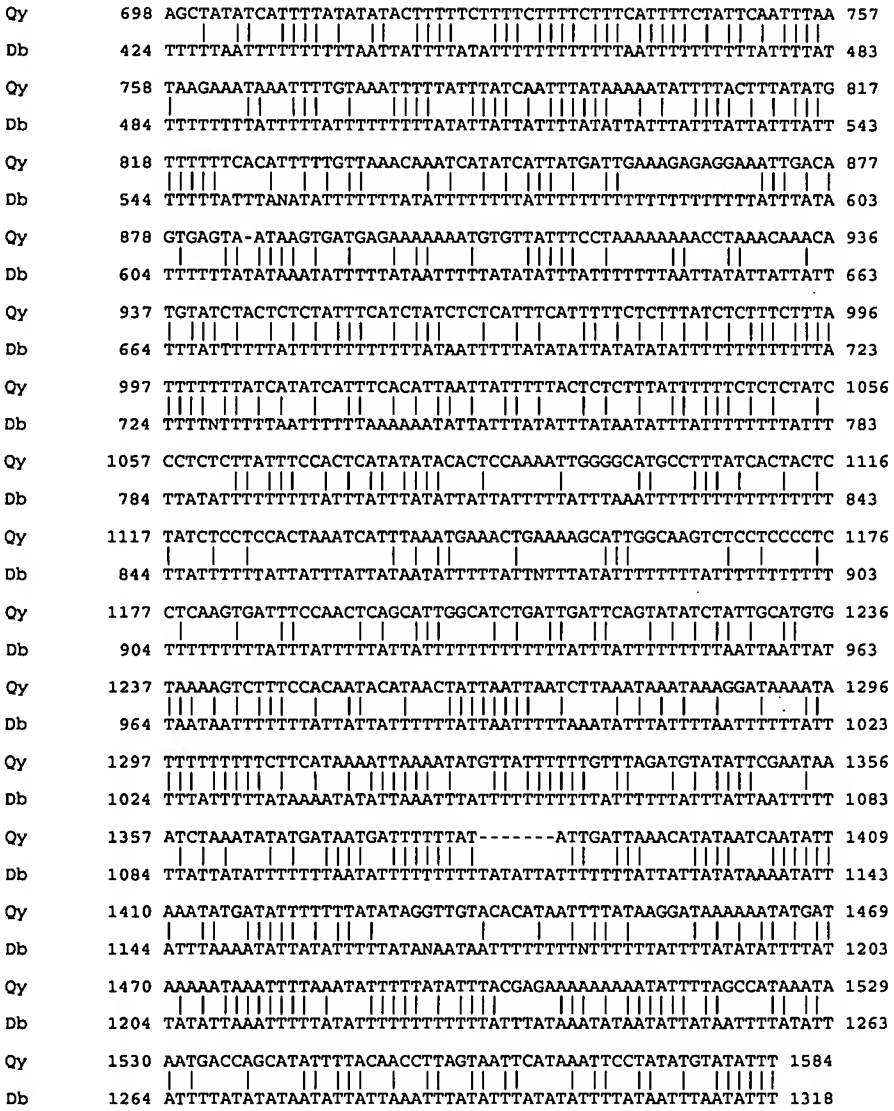
RESULT 1
CF238805/c
LOCUS CF238805 1626 bp mRNA linear EST 05-AUG-2003
DEFINITION AGENCOURT_15099447 NICHD_XGC_Emb6 Xenopus tropicalis cDNA clone IMAGE:6995950 5', mRNA sequence.
ACCESSION CF238805
VERSION CF238805.1 GI:33442013
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1626)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14680 row: g column: 21
High quality sequence start: 71
High quality sequence stop: 316.
FEATURES Location/Qualifiers
source 1. .1626
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6995950"
/tissue_type="neurula"
/dev_stage="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_XGC_Emb6"
/note="Vector: pCMV-SPORT6.1; Site_1: NotI; Site_2: EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.1 kb. Constructed by Invitrogen. Note: This is a Xenopus Gene Collection (XGC) library."
ORIGIN

```
  Query Match             8.6%;  Score 172.4;  DB 7;  Length 1626;
  Best Local Similarity   45.8%;  Pred. No. 2e-14;
  Matches  526;  Conservative    0;  Mismatches  617;  Indels    5;  Gaps    3;

Qy        495 TATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTAT 554
              |||| ||| | || ||  ||   | | || |   || |   |||| || |  ||  ||||
Db       1174 TATTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTAATTTTTTTTTTAT 1115

Qy        555 TATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGT 614
              | | ||  |||||| | | || | ||   | | |||| |  || || |||     | | |
Db       1114 TTTTTTTTTTTTTTTTTTTTTTTTTTAATTTTTTTTTATTTTTTTTTTTATTTTTTTT 1055

Qy        615 AAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGA 674
              |   || | |||||  | | |  ||      ||| |   || |||         |
Db       1054 TTTTTTTTTTTTTTTTTTATTTTTTTTTTNATTTTTTTTTTTTTTATTTTTTTTTTT 995

Qy        675 TATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTT 734
              |||   |  ||| || | |||    | | |  |||||| | |  ||| | |||| ||||
Db        994 TATATTATTTTTTTTTTATTTTTTTTTTTTTTTTTTTATTTTTTTTTTNTTTTTTTTTT 935

Qy        735 CTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTT 794
              |||  |||| | || |   ||  |       |   |||| |   |||||||| | |||||
Db        934 TTTTTTTTTTTTTTTAATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTAAATTT 875

Qy        795 ATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTAT 854
              ||  | ||||    | ||| ||||||    ||||  |         |   |  | | |
Db        874 ---AATTTTTTTTNTNATTTATTTTTTTTTNTTTTTTANTNNTTTTTTTTTAAAAAATTT 818

Qy        855 GATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTC 914
              ||              ||     |     |   | |              | | || |||
Db        817 TTTTTTTTTTTTTTTTTTTNNTTTTTNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTA 758

Qy        915 CTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATT 974
              |    | |||  |          | | | |  | | |||   |    | |    | | |||
Db        757 ATTNNATAAAATTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTACTTTAATTATTATT 698

Qy        975 TTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACT 1034
              ||| | ||| |    ||  || |||||| | |  | |  ||     ||  || |||||  |
Db        697 TTTTTTTTTTTNANATTTNTTNTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTT 638

Qy       1035 CTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTG 1094
              | ||| |||||      ||   | | ||||||   |  | | |    |      |
Db        637 TTTTTTTTTTTTANNANANATAATTATATTATTTTTTTTTTTTTTTTTTTTNNNNNANNN 578

Qy       1095 GGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAG 1154
              |      ||   |  | | |  |     |  |           |||   ||||
Db        577 NAAAAAAATATTTATAAATTTTTTTTTTTTTTTTTTTNTNNNNNNNNNNAAAAAAAAAAT 518

Qy       1155 CATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGA 1214
              |      | |  |    | |             ||   |  |          |   |
Db        517 TAAAATTTTTTTTTTTTTTTTTTTTTNNNNANANAAAAAAAAAAAAAAAANAAAAAAAAA 458

Qy       1215 TTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCT 1274
              |   | | | | ||   | |   |||        || | | ||  |  | ||  |
Db        457 ATTTATNTTTTTTTTTTTTTTTAANAAAAANAAAAAAANAAANAAANAAAAAAAATTTTTTA 398

Qy       1275 TAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTT 1334
              || |||   |  |      | ||||||||||||  |   || |  || |    |  ||||||
Db        397 TATATATCTCTATATATATATATTTTTTTTTNTNTNTATATATANATATNTNTTTTTTTT 338

Qy       1335 GTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAA 1394
              ||    | | |||     |||  |   | | ||||    ||   | | |||||  ||  |
Db        337 TTTNTNTTTTTTATATATATANANANATANATATATATATATATATATATATATATAT 278

Qy       1395 CATATAATCAATATTAAATATGATATTTTTTTATATAGGT-TGTACACATAATTTTATAA 1453
              ||  ||||  |     ||||  | | ||||||  | | || | |||  | ||||
Db        277 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 218

Qy       1454 GGATAAAAAATATGATAAAAATAAATTT-TAAATATTTTTATATTTACGAGAAAAAAAAA 1512
              |||  | |     ||| | ||| || | || |||| | ||||| ||     | | | |
Db        217 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 158
```

```
Qy       1513 TATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATTCCT 1572
              | ||   ||| ||| ||     |       | || |    ||   ||  ||| ||   |
Db        157 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 98

Qy       1573 ATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGAAGGAATCCTACGTCATCTC 1632
              |||| ||||| |  |  || | | | |       ||    |   | |   || ||
Db         97 ATATATATATATATATATATANANANANANNANGNGNGGCGNGNNNCCATTNNTCCTCGG 38

Qy       1633 TTGCCATT 1640
                 |||
Db         37 GANNNATT 30
```

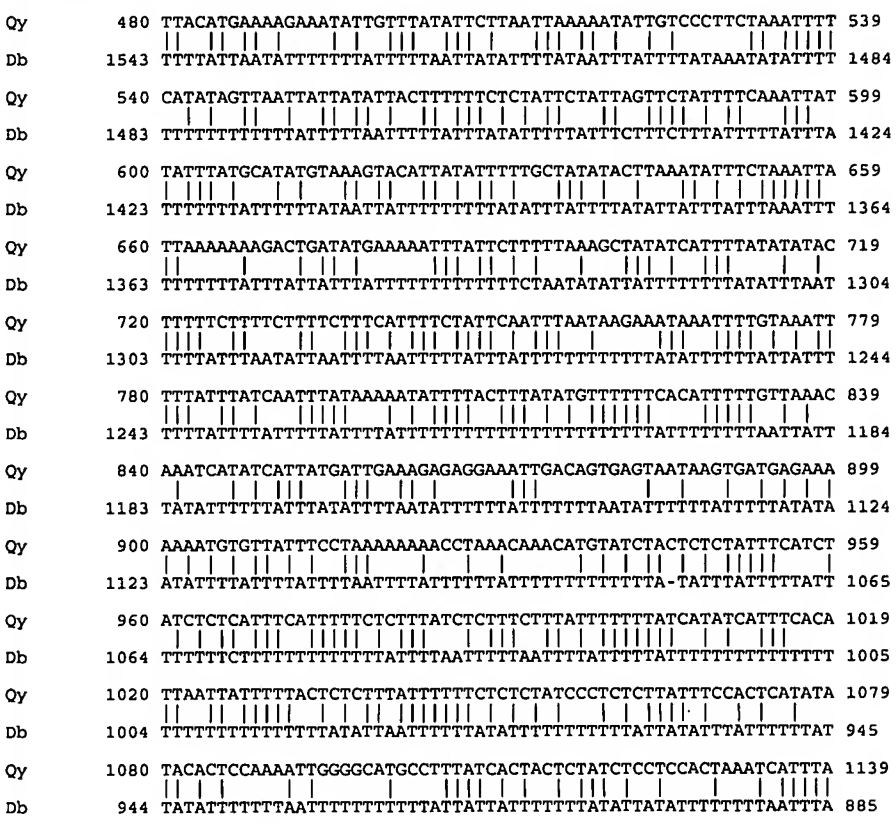
```
RESULT 2
CG749499
LOCUS       CG749499                1348 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P043-4-A06.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG749499
VERSION     CG749499.1  GI:37970425
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1348)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .1348
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN
```

```
  Query Match             8.2%;  Score 165;  DB 9;  Length 1348;
  Best Local Similarity   48.4%;  Pred. No. 2.2e-13;
  Matches  549;  Conservative    0;  Mismatches  571;  Indels   15;  Gaps    3;

Qy        465 TACATCCGTCCCTCATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTG 524
              || ||   |   |  ||| |  ||       |||| ||  ||   ||  ||     |||
Db        184 TAAATATTTTTTTTTTTATAAAAATTATTTTATTATTATTAAATTTTTTTNTTTTTATAA 243

Qy        525 TCCCTTCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTT 584
              |   ||   || |||  |||  |   || || | ||  || ||| | ||||   ||| ||
Db        244 TTTTTTTATAAATTTTTTATTATATTTTTTTTTTTTTTTATTTTTATATTTATTTATTT 303

Qy        585 CTATTTTCAAATTATTATTTATGCATATGTAAAGTACA-------TTATATTTTTGCTAT 637
              | ||||  | |||| |||||    || | || | |          ||||  |||||  | |
Db        304 TTTTTTTTTATTTATAATTTAATTATTTTTATATTTTTTTTATTTTTATTTTTTTTTTTT 363

Qy        638 ATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAA 697
              ||  ||   | |||  ||  |  || |  |     |  | | |   |||  | |||||
Db        364 ATTTTTTTTTTTTTTAAATATTATATATTATTTATTTTTTTTATTTTTTTATTTTTTTTT 423
```



RESULT 3
CL118721/c
LOCUS CL118721 1608 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-72J8_T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-72J8, genomic survey sequence.
ACCESSION CL118721
VERSION CL118721.1 GI:40612356
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1608)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGGG
Class: BAC ends
High quality sequence start: 252
High quality sequence stop: 345.
FEATURES Location/Qualifiers
source 1. .1608
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISB1-72J8"
/clone_lib="ISB1"
/note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC Library Segment 1"
ORIGIN

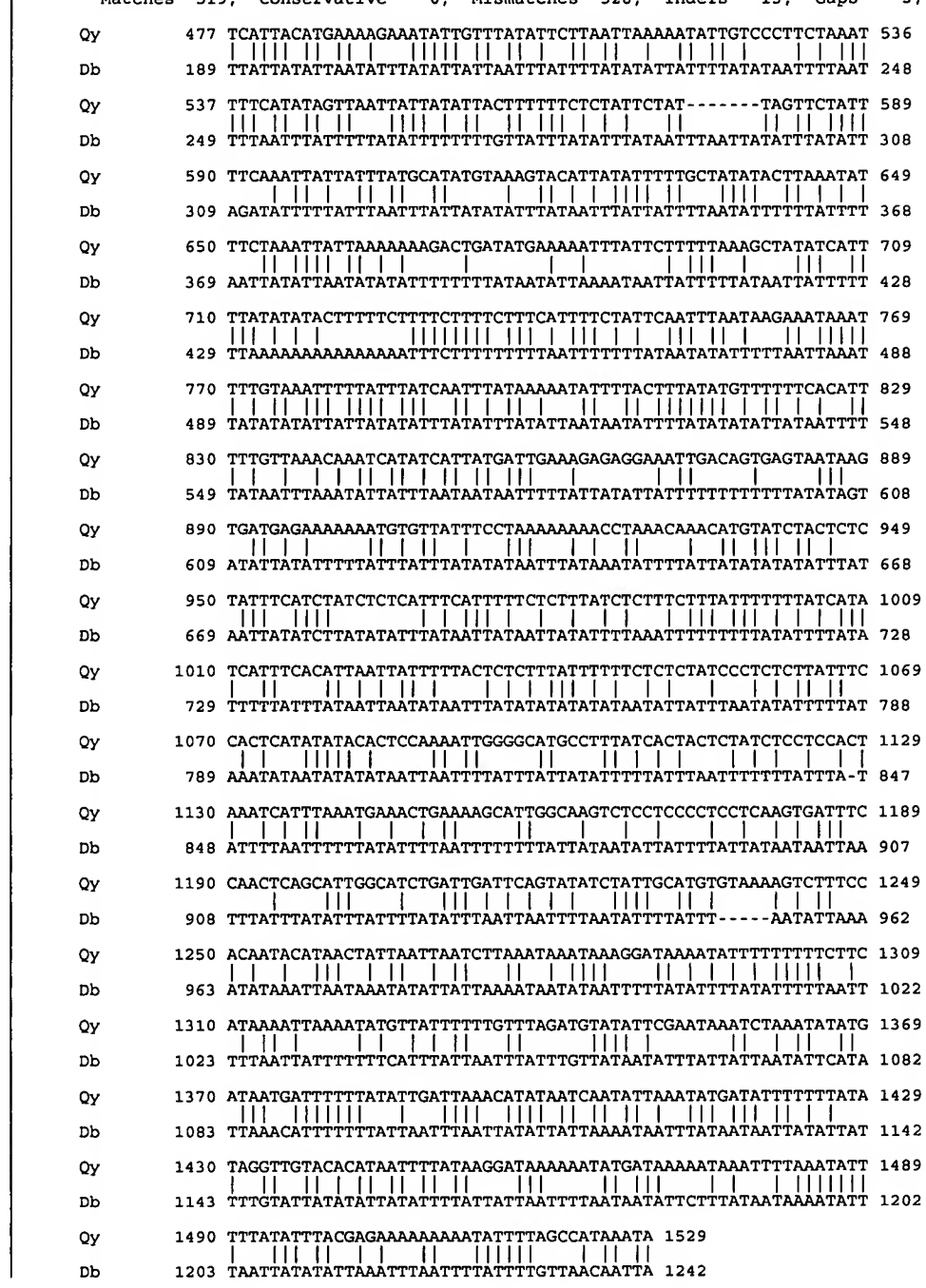
Query Match 7.9%; Score 159.8; DB 9; Length 1608;
Best Local Similarity 48.6%; Pred. No. 1.1e-12;
Matches 503; Conservative 0; Mismatches 522; Indels 10; Gaps 2;



RESULT 4
CL068807
LOCUS CL068807 1242 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-115B3_Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-115B3, genomic survey sequence.
ACCESSION CL068807
VERSION CL068807.1 GI:40524720
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1242)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 7
High quality sequence stop: 57.
FEATURES Location/Qualifiers
source 1. .1242
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-115B3"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"
ORIGIN

Query Match 7.9%; Score 159.2; DB 9; Length 1242;
Best Local Similarity 49.0%; Pred. No. 1.4e-12;

Matches 519; Conservative 0; Mismatches 528; Indels 13; Gaps 3;



```
RESULT 5
AG386981
LOCUS       AG386981               1542 bp    DNA     linear   GSS 03-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-201G10.TJ, genomic survey
            sequence.
ACCESSION   AG386981
VERSION     AG386981.1  GI:47998186
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1542)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1542
                     /organism="Mus musculus molossinus"
                     /mol_type="genomic DNA"
                     /sub_species="molossinus"
                     /db_xref="taxon:57486"
                     /clone="MSMg01-201G10.TJ"
                     /sex="male"
                     /tissue_type="mixture of kidney and spleen"
                     /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN

  Query Match             7.8%;  Score 156;  DB 9;  Length 1542;
  Best Local Similarity   47.8%;  Pred. No. 3.8e-12;
  Matches  507;  Conservative    0;  Mismatches  534;  Indels   19;  Gaps    2;

Qy        492 AAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAAT 551
              ||||| | | | | |   ||  || |  |    |   ||     | || |  || |||||
Db        299 AAATAATTTATTTTTATATATATATATTTTAATTTNTTAATTTTATTAAATTATTTAAT 358

Qy        552 TATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATA 611
              | ||| |||  ||||| | | |  |||   || || | | || || || ||||
Db        359 AAATATTTTATATTTTTTTTTTTTATTATATATTTTAATATAAATTTTTTTTTTAATATAT 418

Qy        612 TGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGAC 671
              | || |   || |  |||   | |  | ||||||| |   ||||    | ||| |
Db        419 TTTATAAATTTTTTTTATTTATAATTTTTATAAAATATTTTTTTTTTATATTATAAATAT 478

Qy        672 TGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCT 731
                |||  |    |||||  |||||| | |     ||||    | || | ||| || | |
Db        479 AAATAATATTTTTTTATATTTTTTATATATTATAAATTTATATTTTATTATTTTTTATAT 538

Qy        732 TTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAA 791
              ||| | |  | |  | ||  || |||||     | | ||||   ||||    || || |
```

```
Db        539 TTTTTATATTATATTTTTATATATAATATATTTTTATTTTTAATAATTANATTTAATAAT 598

Qy        792 TTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCAT 851
              ||  || | | | ||| ||| ||| || |||    || ||   ||    ||    | |
Db        599 ATTTAAATATTTTATTATTTTTTATTTTATTTATTTTTTATTTAAAATATTATTTATTATT 658

Qy        852 TATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTAT 911
              | |  ||  ||         |||   | |   | ||   |  |   |  |   | |  |||
Db        659 TTTATTTATAAATATATTTATTATTATTTTTTTATTTTTATTTTTATTATTTTTTTATAT 718

Qy        912 TTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTC 971
              ||  | |  |  |  |  |        ||  || || | | | ||   |  || | |  ||
Db        719 TTTTTTATTATTATATTTTATTTTATTAATTTATTTTTTTTTATTTTNATATTTTTATTT 778

Qy        972 ATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTT 1031
              |||||  | |||  |   ||| |||  || | | |  |     |      |  | |||| ||
Db        779 TTTTTTATTTTTTTTTATTTTTTTTTATTATATTTATTTAATTAATTTTTATAATATTATT 838

Qy       1032 ACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAA 1091
              | |   | | ||  || |  |          ||||  |  |   ||| || | |  |
Db        839 ATTTATTATTTTAATTATAATTTAAATATATTTATATATATATTTATTTATAATTTATTT 898

Qy       1092 TTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAA 1151
              ||      |    |||     |    ||  |       |  ||  | || || |   | |
Db        899 TTTTTTTAATAATTTTATTATTTATTAATTATAATTNTTTATTTTATATATTATTATTTA 958

Qy       1152 AAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGAT 1211
                  |  | |  |     | |  |  |||| |  |     || |     ||   | | |
Db        959 TTAATATTATATATATATTATTATATTATNTTTATTTTTAAATATTATTTTTTTTATATT 1018

Qy       1212 TGATTCAGTATATCTATTGCATGTGTA------AAAGTCTTTCCACAATACATAACTATT 1265
              | |||   | |||  | ||     | ||      |   | |||    | || ||   |  |
Db       1019 TTATTTTTTTTATTTNTTATTATTTTATTAAATANTTTATTTATTAATTATATTTTTTAT 1078

Qy       1266 AATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATA------------ 1312
              | |   | |||  |||| |     ||  | |||| |||||| || |||
Db       1079 ATTATTTATTATATAATTTTCTTATTTANTATTATTTTTTTTTTTATATTTAATATATTTT 1138

Qy       1313 AAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATA 1372
              |   | ||  |||  |||||| |  |||  | | | |    |   ||| || |||||
Db       1139 ATTTAAATTTATANTATTTTATTATTAATTTTTTTTAATATNTAATATATATATAATTAT 1198

Qy       1373 ATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAG 1432
              || ||||| |   |   ||    | | || | ||||||   | |||||  | ||  ||
Db       1199 ATTATTTTATNATTATTTTTTTTTTTNTTATTATTATTAATATTTATATTATATTTAATTA 1258

Qy       1433 GTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTT 1492
               || |     | |||||||  |   ||||  ||||  |    |||   | || |||| |||
Db       1259 TTTATTTTATTTATTTTATTATTNTAAATTATATATTTTTTATATTATATATATATATTT 1318

Qy       1493 ATATTTACGAGAAAAAAAAAATATTTTAGCCATAAATAAAT 1532
              |  | ||      | | | | |||||     || | | | |
Db       1319 AATTNTAATTATATATATATTATTTATTTAATTATTTACT 1358
```

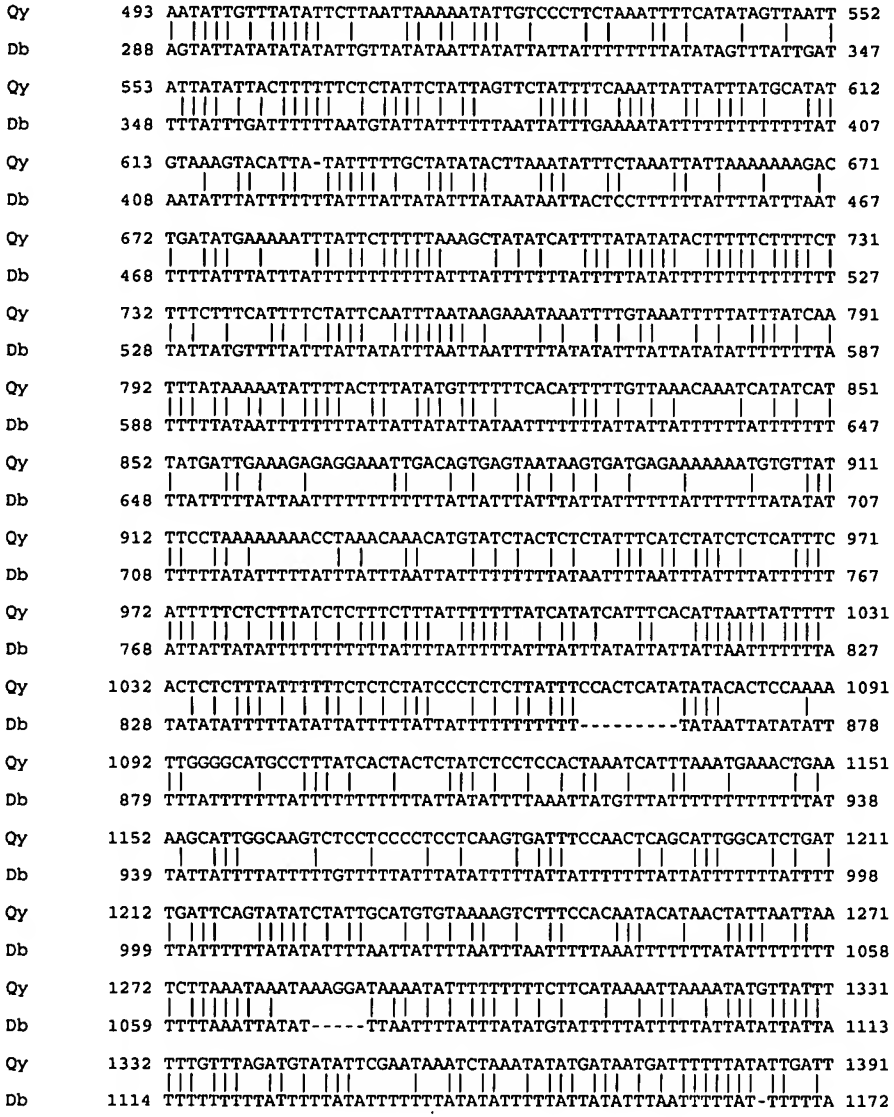
```
RESULT 6
CG753083
LOCUS       CG753083               1896 bp    DNA     linear   GSS 24-OCT-2003
DEFINITION  P048-1-C01.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
            genomic survey sequence.
ACCESSION   CG753083
VERSION     CG753083.1  GI:37977199
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 1896)
  AUTHORS   Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K.,
            Buntjer,J., van der Meulen,M. and Sommer,R.J.
  TITLE     An integrated physical and genetic map of the nematode Pristionchus
            pacificus
  JOURNAL   Mol. Genet. Genomics 269 (5), 715-722 (2003)
```

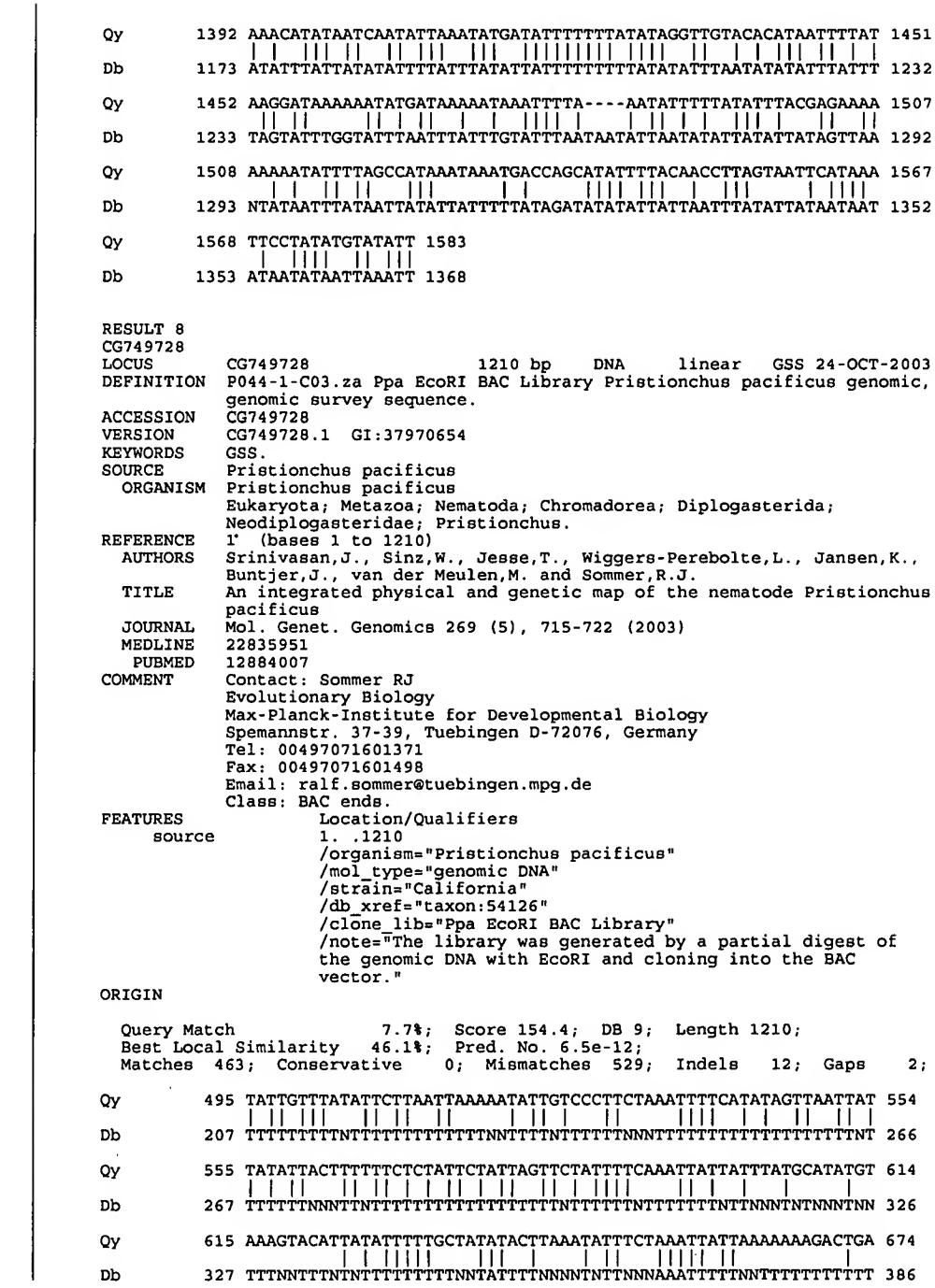
```
  MEDLINE   22835951
   PUBMED   12884007
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 00497071601371
            Fax: 00497071601498
            Email: ralf.sommer@tuebingen.mpg.de
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .1896
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /db_xref="taxon:54126"
                     /clone_lib="Ppa EcoRI BAC Library"
                     /note="The library was generated by a partial digest of
                     the genomic DNA with EcoRI and cloning into the BAC
                     vector."
ORIGIN

  Query Match             7.7%;  Score 155.8;  DB 9;  Length 1896;
  Best Local Similarity   48.6%;  Pred. No. 3.9e-12;
  Matches  574;  Conservative    0;  Mismatches  584;  Indels   24;  Gaps    6;

Qy        479 ATTACATGAAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTT 538
              |||| ||      || | || ||| || | ||| |   | | |||     | || |||
Db        694 ATTATATATTTTTAATTTTTATTTTTAATTTTATTATTATAAATTTATTAATTTATATTA 753

Qy        539 TCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTA 598
              | ||||| || | || ||| | | | ||        |   |  || ||  |||||  ||||
Db        754 TTATATAATTTAATAATATTTAAATATTAAAAAATTATTTATATTTTAATTTTATAATTT 813

Qy        599 TTATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATT 658
              ||   |    |  | || |  |   ||   || |  | |     ||  || ||| |   ||
Db        814 TTTNATNATAAATTATATATAAATATAATATTATTATTTTATTTTNTATTTTTTTTTTTT 873

Qy        659 ATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATA 718
              | |  |  |  | |  |   |  |   || | |||| | |  |  ||  |||| |  ||
Db        874 AATTTATTTATTTTTTTTATATTATAATAATATTTTAATAATTTAATTNTTTTTTTAATT 933

Qy        719 CTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAAT 778
                |   |||  ||||  |  |||||  ||  || |||||||  || | || |  | |
Db        934 AAATAAATTTATTTTTAATATATTTTTAAT--AAATTAATAAATAAAATATNTATTTATA 991

Qy        779 TTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAA 838
              |||  || |  | ||  || |||| ||||| | |||||  | |             ||||
Db        992 TTTATTTAAAAATTTATTATAAATTTTTTATTATATATAATAT-----------ATAAA 1039

Qy        839 CAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAA 898
               ||||   || | ||| ||  ||| | |  ||  | | |      ||||||  | |  | |
Db       1040 NAAATATAATTAATATTATATAAATAAATAAATATTATATATTTTAATAATTAAATAAAT 1099

Qy        899 AAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTAT-TTCAT 957
              | |  | | || |||  |  | || |  ||         | ||| || | | | | || ||
Db       1100 ANATTTTTATTTTTTAATTTATAATATTTATTTTTTTTTATATTTATTTTTTTTATTTAT 1159

Qy        958 CTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCA 1017
              |   | |  ||  |  ||  | |||||     || ||||||||||     |    |||
Db       1160 ATTATTATATTTATAAATTAATTTTTATTAAATTATTTATTTTTTATATAATATATTAAT 1219

Qy       1018 CATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATA 1077
              ||  |||||||||  | |  |||||||| |         |   |||| |  | | |
Db       1220 TATATATTATTTTTTTTTTTATTATTTTTTTTNATATATTATTATTTATATAAATTTTTT 1279

Qy       1078 TATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATT 1137
              | |    |   ||         |  ||||  |    |  |          | | ||| |||
Db       1280 TTTTTTTT---ATTTTTATAATTTAATTTAATAAATTTAAAAAAATAATATTTAATTATT 1336

Qy       1138 TAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAG 1197
              | ||| |   |          |                    | || |   ||  || |
Db       1337 TTAATTATNATATTTTTATATATATTTATATTNNTNAAATTTAAATATATTTTAAATATT 1396

Qy       1198 CATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACA 1257
              | |   ||   ||| | | |  || | | ||   |   |||   | |||  | |    |
Db       1397 TAATAAAATAAAATTTANTTATAATTTTTTTTTTATTTNATAATTTTATTTATAAATATAA 1456

Qy       1258 TAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATT 1317
              ||  |   |  ||   | | ||| |||    |  ||| | | | |  |  |  || |
Db       1457 TATTTTANATAAAAAAAAATAAAAAAAAATATAATTTTATNTATATATTATTTTTTAAAT 1516

Qy       1318 AAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGAT 1377
              |||||||   |  | || | || ||  | | |  ||  |    ||||||||| ||||| ||
Db       1517 AAAATATTAAAAATATTATATATATTAAAAATAAAAATATAAAAAATATATAATAATTAT 1576

Qy       1378 TTTTTATATTGATTAAACATATAATCA----ATATTAAATATGATATTTTTTTATATAGG 1433
              || | ||  |  || ||   || || |    ||||  ||  | | | |  ||  | |
Db       1577 TTATAATNATATTTTAAATAATTATTATTATATATATAAATTNAAAATAATTATTTTTTT 1636

Qy       1434 TTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTA 1493
              |  ||   ||| |||  | |   | ||  ||||  || | ||||  |||| || ||| ||
Db       1637 TNNTAATTATATTTTATTTATTTTTAATTATATTNTATATATAA--TTTANATTTTTATA 1694

Qy       1494 TATTTACGAGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCT 1553
              | | ||  |  |   || |||||       ||  ||| ||        |  | |||  |  |
Db       1695 TTTATATNAATATTTAATATATTAATAATATTTATATATTTATTTAAATTNTTATTATTT 1754

Qy       1554 TAGTAATTCATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAAGGGA 1613
                || || ||   ||  | | |   | ||| |   || | | |  || |  |||    |
Db       1755 ATATATTTAATTTTTTATTTTTTAATTTTTTTATTATATATATATTTATTTTTTATTANA 1814

Qy       1614 AGGAATCCTACGTCATCTCTTGCCATTTGTTTTTCATGCAAA 1655
                |  ||  | | | |     || |||||     |||
Db       1815 TTTTTTTTTATATTANATTTAATTTATTATTTTTNTAATAAA 1856


RESULT 7
AG350209
LOCUS       AG350209               1378 bp    DNA     linear   GSS 02-JUN-2004
DEFINITION  Mus musculus molossinus DNA, clone:MSMg01-146M02.TJ, genomic survey
            sequence.
ACCESSION   AG350209
VERSION     AG350209.1  GI:47923519
KEYWORDS    GSS.
SOURCE      Mus musculus molossinus
  ORGANISM  Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     BAC end Sequences of Library MSMg01
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1378)
  AUTHORS   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the mouse BAC library MSMg01. For BAC
            library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
            Tsukuba Institude, Bio Resource Center,
            The Institute of Physical and Chemical Research (RIKEN) 3-1-1
            Koyadai, Tsukuba, 305-0074 Japan
            phone: 81-298-36-9189, fax: 81-298-36-9199
            e-mail: abe@rtc.riken.jp
            PRIMERS
            Sequencing : TJ
            LIBRARY
            Vector     : pBACe3.6
            R.Site 1   : EcoRI
            R.Site 2   : EcoRI.
FEATURES             Location/Qualifiers
     source          1. .1378
```

/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-146M02.TJ"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

ORIGIN

Query Match 7.7%; Score 155.4; DB 9; Length 1378;
Best Local Similarity 49.6%; Pred. No. 4.6e-12;
Matches 544; Conservative 0; Mismatches 532; Indels 20; Gaps 5;

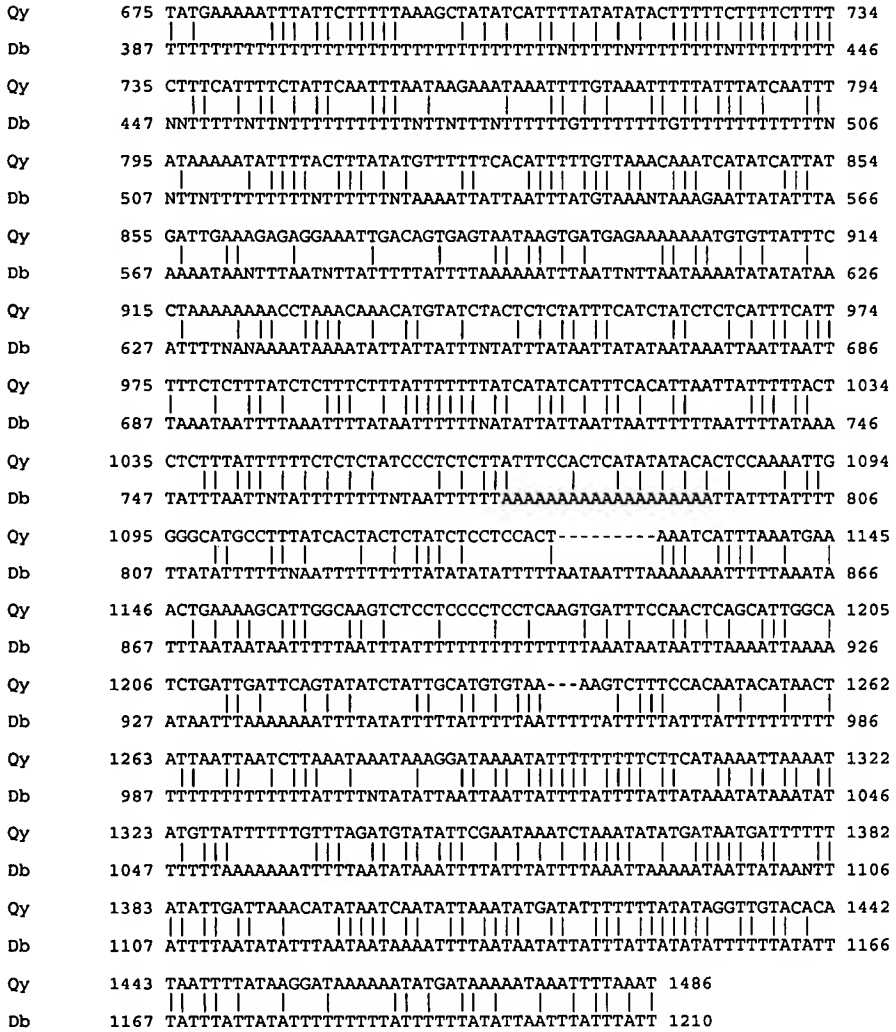




RESULT 8
CG749728
LOCUS CG749728 1210 bp DNA linear GSS 24-OCT-2003
DEFINITION P044-1-C03.za Ppa EcoRI BAC Library Pristionchus pacificus genomic, genomic survey sequence.
ACCESSION CG749728
VERSION CG749728.1 GI:37970654
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1210)
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K., Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .1210
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."

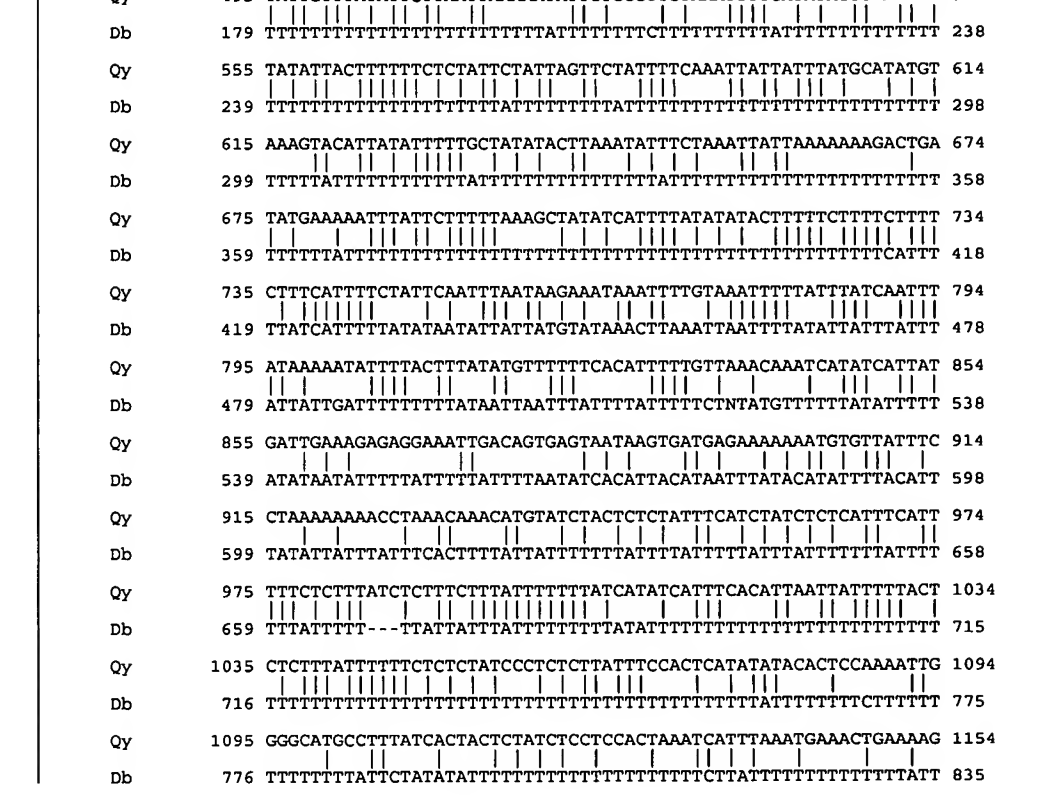
ORIGIN

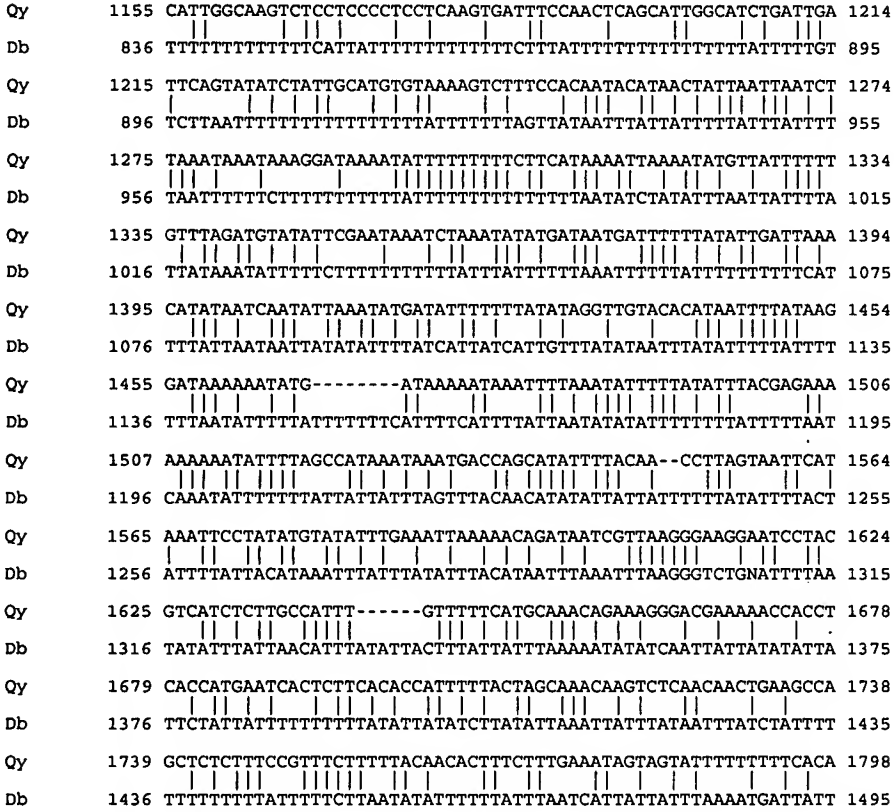
Query Match 7.7%; Score 154.4; DB 9; Length 1210;
Best Local Similarity 46.1%; Pred. No. 6.5e-12;
Matches 463; Conservative 0; Mismatches 529; Indels 12; Gaps 2;



RESULT 9
CL080711
LOCUS CL080711 1805 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-159I8_RM4.1 CH216 Xenopus tropicalis genomic clone CH216-159I8, genomic survey sequence.
ACCESSION CL080711
VERSION CL080711.1 GI:40536624
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1805)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM4 ctcaagggcatcggtcgagc
Class: BAC ends
High quality sequence start: 690
High quality sequence stop: 766.
FEATURES Location/Qualifiers
source 1. .1805
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-159I8"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"
ORIGIN

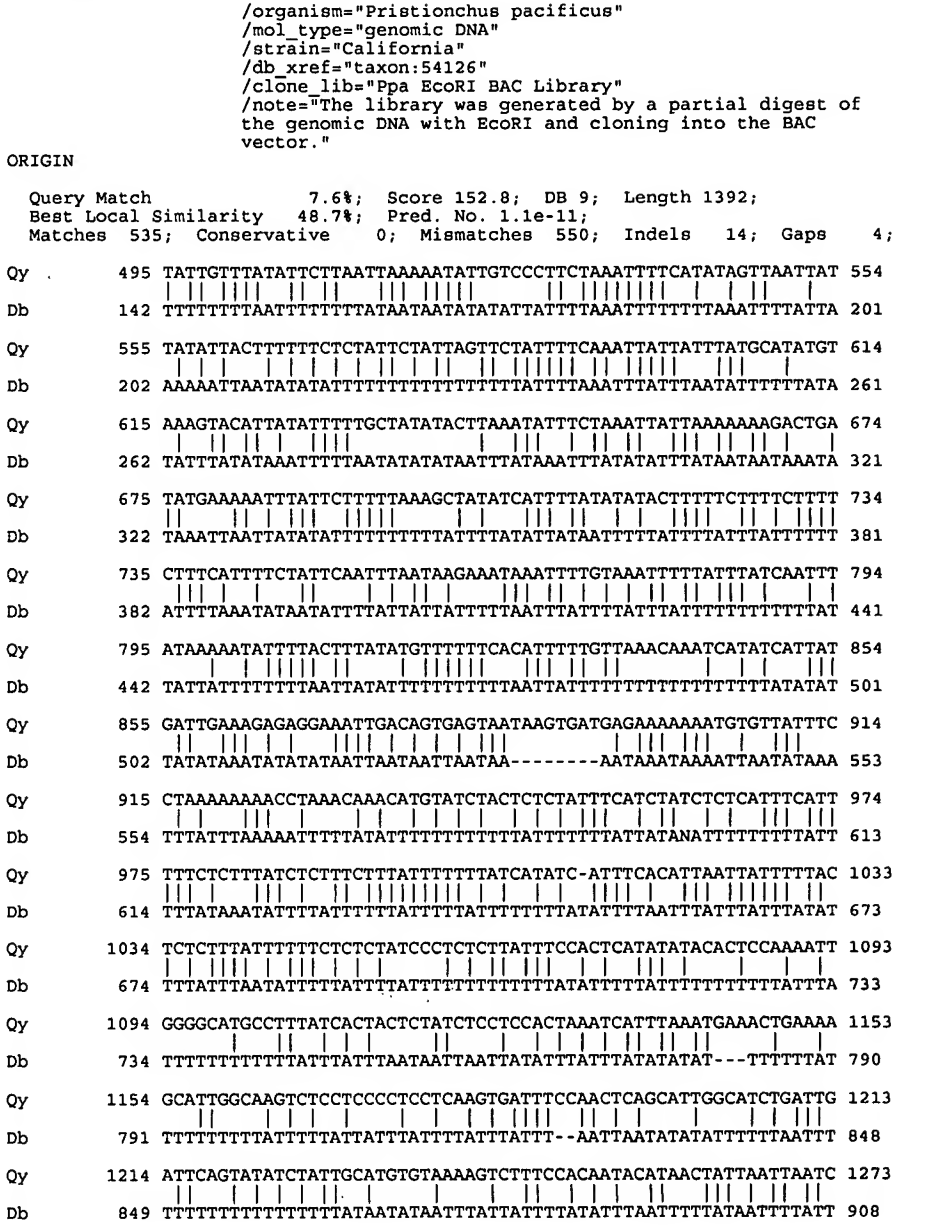
Query Match 7.6%; Score 153; DB 9; Length 1805;
Best Local Similarity 46.9%; Pred. No. 9.5e-12;
Matches 623; Conservative 0; Mismatches 687; Indels 19; Gaps 4;

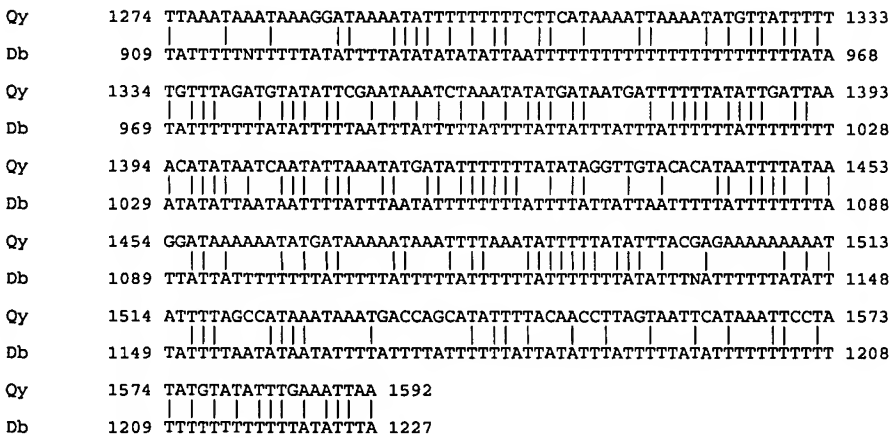




RESULT 10
CG757503
LOCUS CG757503 1392 bp DNA linear GSS 24-OCT-2003
DEFINITION P052-4-C08.za Ppa EcoRI BAC Library Pristionchus pacificus genomic, genomic survey sequence.
ACCESSION CG757503
VERSION CG757503.1 GI:37986131
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1392)
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Perebolte,L., Jansen,K., Buntjer,J., van der Meulen,M. and Sommer,R.J.
TITLE An integrated physical and genetic map of the nematode Pristionchus pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .1392
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Ppa EcoRI BAC Library"
/note="The library was generated by a partial digest of the genomic DNA with EcoRI and cloning into the BAC vector."
ORIGIN

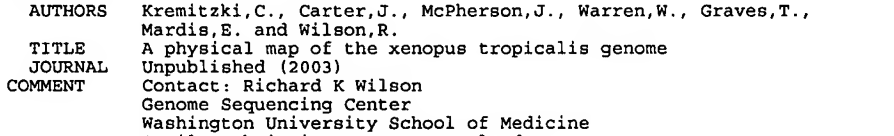
Query Match 7.6%; Score 152.8; DB 9; Length 1392;
Best Local Similarity 48.7%; Pred. No. 1.1e-11;
Matches 535; Conservative 0; Mismatches 550; Indels 14; Gaps 4;

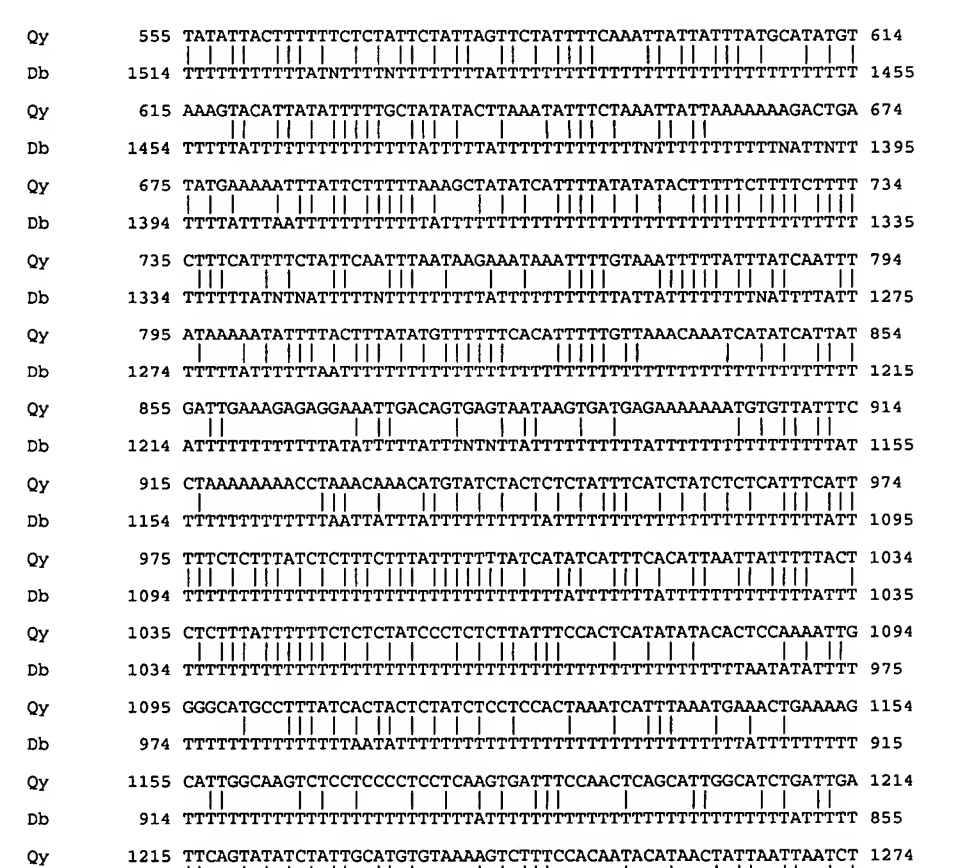




RESULT 11
CL082569/c
LOCUS CL082569 1632 bp DNA linear GSS 31-DEC-2003
DEFINITION CH216-167P5_Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-167P5, genomic survey sequence.
ACCESSION CL082569
VERSION CL082569.1 GI:40538482
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus; Silurana.
REFERENCE 1 (bases 1 to 1632)
AUTHORS Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T., Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: Sp6 ATTTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 1062
High quality sequence stop: 1114.
FEATURES Location/Qualifiers
source 1. .1632
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-167P5"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis BAC library"
ORIGIN

Query Match 7.6%; Score 152.4; DB 9; Length 1632;
Best Local Similarity 47.2%; Pred. No. 1.2e-11;
Matches 474; Conservative 0; Mismatches 525; Indels 5; Gaps 1;





RESULT 12
CL110653
LOCUS CL110653 1594 bp DNA linear GSS 05-JAN-2004
DEFINITION ISB1-53P23_Sp6.1 ISB1 Xenopus tropicalis genomic clone ISB1-53P23, genomic survey sequence.
ACCESSION CL110653
VERSION CL110653.1 GI:40604288

```
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 75000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 390
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .1594
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:8364"
                     /clone="ISB1-53P23"
                     /clone_lib="ISB1"
                     /note="Vector: pBeloBAC11; ISB-1 Xenopus tropicalis BAC
                     Library Segment 1"
ORIGIN
```

```
  Query Match             7.6%;  Score 152;  DB 9;  Length 1594;
  Best Local Similarity   47.7%;  Pred. No. 1.3e-11;
  Matches  479;  Conservative    0;  Mismatches  520;  Indels    5;  Gaps    1;

Qy        495 TATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTAT 554
              | || ||||| || ||| ||      | || |   || |   ||||  ||   || ||| |
Db        313 TTTTTTTATTTTATTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAATTTTTATTTT 372

Qy        555 TATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGT 614
              | | ||| | |||| | | || | ||  || | ||||    || || ||| |   | | |
Db        373 TTTTTTATTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTT 432

Qy        615 AAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGA 674
                |   || | |||||  | | |  ||   |||||  ||  || || |   |  | | |
Db        433 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTATTTTTTTATTTATTATTTA 492

Qy        675 TATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTT 734
              | |       | | || |||||     | | |  |||| | |||| ||||| |||| ||||
Db        493 TTTTTTTTTATTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTATATTTTTTTTTTTTTTT 552

Qy        735 CTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTT 794
              |||  | || | ||   |||| |       |   |||| |   |||||  ||  |  || | |
Db        553 TTTTTTTCTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTATAT 612

Qy        795 ATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTAT 854
                ||   || ||||  | | ||| ||||||    |||||  ||        |  | | ||| |
Db        613 TTATTTATTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTT 672

Qy        855 GATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTC 914
                ||                ||      |   ||      |  |           | | || |||
Db        673 TTTTTTTTTTTTATTTTT-----TTTTTATATTTTTTTTTTTTTTTTTTTTTTTTTTTT 727

Qy        915 CTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATT 974
              |  | |      |          | ||| |  ||| ||| |    |   | |  |||  ||
Db        728 TTTTATATTTTTTTTTTTTTATTTTATTTTTTCTTTATATTTATTTATTTTTTTTTTTT 787

Qy        975 TTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACT 1034
              | | | ||  | | ||| ||||||||||| |  | | ||||    ||  |||||| |  |
Db        788 TATTTTTTATTTTTTTTTTTATTTTTTTTTTTTTTTATTTTTTTTTTTTTATTTATTTT 847

Qy       1035 CTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTG 1094
                | ||| ||||||  | |   |   | | || |||    | || |||    |   |  ||
```

```
Db        848 TTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTATTTATTTGTTTGTATTTTT 907

Qy       1095 GGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAG 1154
                |   ||| |   || |   | |  |  | | | |  |||    |      |      |
Db        908 TTTTTTTTTTTTTTTTTATTTATTTTTTTTTATTTATTTTTTTTTTTTGTTTTTTTAA 967

Qy       1155 CATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGA 1214
                ||    | | |  |         |    |  |||   | |     ||    | |  ||
Db        968 TTTTTTTATTTTTTTTTATCTTTTTTTTATTTTTTTATTTTTTTTTTTTTTTTTTTT 1027

Qy       1215 TTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCT 1274
              ||   |   | | ||   | | | |  | |||       |   |    | ||  ||  | |
Db       1028 TTTTTTTATTTTTTTTTTTTTATTTATTTTTTTTTTTTTTTTTTTTTTTTGTTTTTTTT 1087

Qy       1275 TAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTT 1334
                |  | |    ||   | |||||||||  ||  |      ||   | | || || |||
Db       1088 CTTTTGTTTTATTATTATTTTTTTTTTTTTTTTTTTTTTTTTTATATTTTTTTTTTTTATTT 1147

Qy       1335 GTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAA 1394
                | |   | | ||||    |   | |   | ||| ||  |   || | | |||| |||
Db       1148 ATGTTTTTTTTTATTGTTTTTTTTTTTTTTTTTATAATTTTTTTTATATTTATTTATTTTT 1207

Qy       1395 CATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAG 1454
                | | |      ||| | | | | | ||||| |   || |     |     ||| | |
Db       1208 TTTTTTTTTTATTTATTTTTTTTTTATATTTATTTTTTTTTTTTTTTAATTATTTTTTAT 1267

Qy       1455 GATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTT 1498
                ||      ||| ||      |   ||||   | ||||||||||
Db       1268 TTTATTTTTTATTATTTTTTTTTTTTTTTTTTTTTTTTTATATTT 1311
```

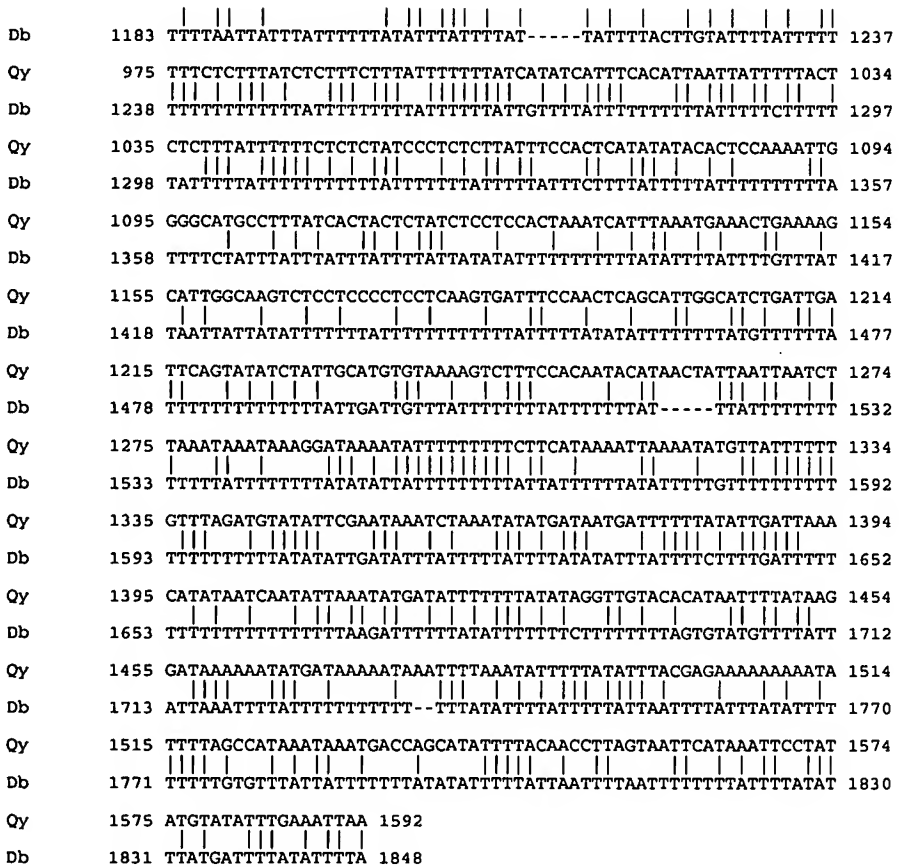
```
RESULT 13
CL038406/c
LOCUS       CL038406                1594 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-46A6_Sp6.1 CH216 Xenopus tropicalis genomic clone CH216-46A6,
            genomic survey sequence.
ACCESSION   CL038406
VERSION     CL038406.1  GI:40494319
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1594)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: Sp6 ATTTAGGTGACACTATAG
            Class: BAC ends
            High quality sequence start: 471
            High quality sequence stop: 522.
FEATURES             Location/Qualifiers
     source          1. .1594
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-46A6"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN
```

```
  Query Match             7.5%;  Score 150.6;  DB 9;  Length 1594;
```

```
  Best Local Similarity   47.9%;  Pred. No. 2.1e-11;
  Matches  481;  Conservative    0;  Mismatches  518;  Indels    5;  Gaps    2;

Qy        495 TATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTAT 554
              | || ||| | || ||| ||  || |||| |   || |   | ||  | | ||  || |
Db       1558 TTTTTTTTTTTTTATTATTTTAATTATTTTTTTTTATTTTNTTNTTTTTTTTTTTTTT 1499

Qy        555 TATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGT 614
              | | ||  |||||| | | ||   ||  || | |||| |   | || ||| |  || | |
Db       1498 TTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTATTATTTTTTTTTTTTATNTAT 1439

Qy        615 AAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGA 674
                |   |||| ||||   ||| |  ||  || | | |   || ||| |  |
Db       1438 TTTTTTATTTATTTTTT---TATTTTTTTTTATTTATTTTTTTTTTTTATATTATTTTNTT 1382

Qy        675 TATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTT 734
              |    |  ||| || ||||       ||| |  |||| | | |  |||||  ||  ||||
Db       1381 TTATTATTTTTTTTTTTTTTATTTTTTATTTTTTTTTTTTTTTTTTTTTTATTNTTTTT 1322

Qy        735 CTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTT 794
                ||  |||  | ||   |||  |   || |   |||| |   ||| ||||| |   |||
Db       1321 TATTTTTTTNATTTTTTTTTTTTTTTTAATTTTTTTTTTTTTTTTTATATTTTTTTTTT 1262

Qy        795 ATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTAT 854
              |    ||||  ||| | | ||||||    ||||| |||        | |  ||||
Db       1261 TTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTATTTTTTTAT 1202

Qy        855 GATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTC 914
              |||   |        ||     |  |   |   |       | | |  |||
Db       1201 TATTTTTATTTTTTTTTTTTATTTTTTTTTTTTATATTTATATTTTTTTTTTATTTTT 1142

Qy        915 CTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATT 974
              | |  |     |  |     || | | || | | | | | || | | | |  |||  ||
Db       1141 TTTATTATTTTTTTTATTTTATTTTTTTATTTTTTTTATTATTTNTTTTTTTTTTTTTT 1082

Qy        975 TTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACT 1034
              ||| | ||||| | ||| ||| ||||||  |  | |  |||    ||  ||||||||  |
Db       1081 TTTTTTTTTATTTTTTTATTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTATTTTTTT 1022

Qy       1035 CTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAATTG 1094
              | ||| ||||||  | | | |  | || |||     |  | ||||    |    |  ||
Db       1021 TTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATATTTTTTTATTTTT 962

Qy       1095 GGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAG 1154
                   ||| |   |  | | | |  |  | |  |   |||    |  |  | | ||
Db        961 TTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTATTTTATAA- 903

Qy       1155 CATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGA 1214
                ||      | |  |   | | |    | |||      |    |||      | |  |
Db        902 -TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTATTTTTTTTTTATTT 844

Qy       1215 TTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAATCT 1274
                ||   | | | ||   | | |      | |||       |   ||  | || |||  | |
Db        843 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTATTTTTT 784

Qy       1275 TAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTT 1334
              |   ||  | |     |   | |||||||| ||  |    ||  | | | ||  |||||
Db        783 TTTTTATTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTATTTTT 724

Qy       1335 GTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAA 1394
                |||      | ||||     |  | |   | ||| ||  |  ||||| | ||  ||  |
Db        723 TTTTTTTNTTTTATTATTTTTTTTTTTTTTTTTATTATTTTTTTTTTTTTTTTTATTTTTA 664

Qy       1395 CATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAG 1454
                | |  |   | ||   | |    |||||||| |||   || |      |  ||||  |
Db        663 TTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTT 604

Qy       1455 GATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTT 1498
                | |       | |||     |   ||||   | ||||| | |||
Db        603 TTTTATTTTATTNATATTTTTTTTTTTTTTTTTTTTTTTTTTT 560


RESULT 14
```

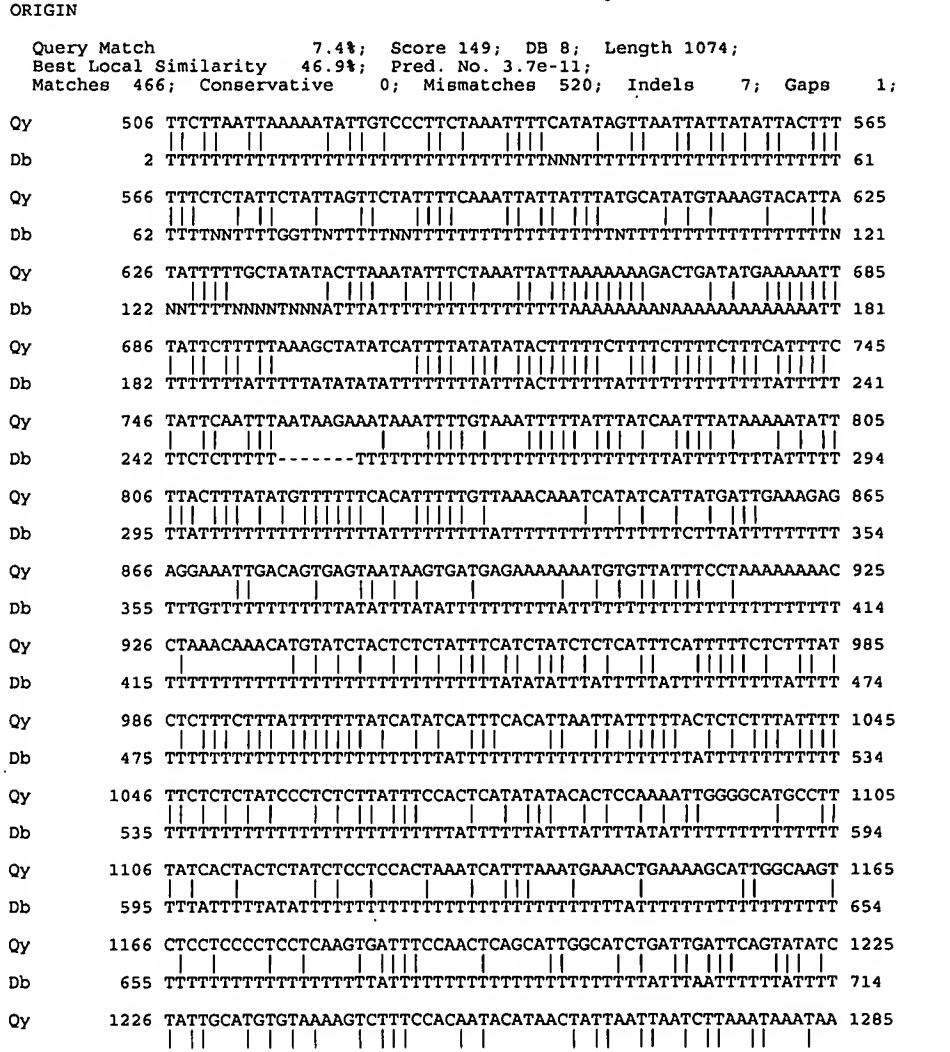
```
CL082000
LOCUS       CL082000                1981 bp    DNA     linear   GSS 31-DEC-2003
DEFINITION  CH216-165P18_RM4.1 CH216 Xenopus tropicalis genomic clone
            CH216-165P18, genomic survey sequence.
ACCESSION   CL082000
VERSION     CL082000.1  GI:40537913
KEYWORDS    GSS.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 1981)
  AUTHORS   Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
            Mardis,E. and Wilson,R.
  TITLE     A physical map of the xenopus tropicalis genome
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Richard K Wilson
            Genome Sequencing Center
            Washington University School of Medicine
            Email: submissions@watson.wustl.edu
            Insert Length: 175000   Std Error: 0.00
            Seq primer: RM4 ctcaagggcatcggtcgagc
            Class: BAC ends
            High quality sequence start: 265
            High quality sequence stop: 497.
FEATURES             Location/Qualifiers
     source          1. .1981
                     /organism="Xenopus tropicalis"
                     /mol_type="genomic DNA"
                     /strain="Nigerian frog"
                     /db_xref="taxon:8364"
                     /clone="CH216-165P18"
                     /sex="male"
                     /cell_line="Stock 248 F7A2, inbred N7"
                     /clone_lib="CH216"
                     /note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
                     BAC library"
ORIGIN
```

```
  Query Match             7.4%;  Score 149.6;  DB 9;  Length 1981;
  Best Local Similarity   48.0%;  Pred. No. 2.8e-11;
  Matches  527;  Conservative    0;  Mismatches  559;  Indels   12;  Gaps    3;

Qy        495 TATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTAT 554
              | || ||| | || | | || |   | || |    ||    | ||  | || ||| |
Db        763 TTTTATTTTTTTTTTAATTTTATTTTTTTATGTTTTATTTTGTTTTTTTATTTTATTTT 822

Qy        555 TATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGT 614
              | | ||  |||||| | | || | ||  || | ||||  |  |  | ||| |  || |||
Db        823 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTATATTCTTTTTTTTTATTTGT 882

Qy        615 AAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGA 674
                |  ||  || |  |||||  | |||  ||   |||||  |   || ||        | |
Db        883 TATTTATTTTTTTATTTTTTTTTTATTATTTTTTATTTTTTTTTTTTTTTTTTTTATTTT 942

Qy        675 TATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTT 734
              | |      |||| || ||| |  |  | |||  | ||||||  |  |||||  |||  ||||
Db        943 TTTTTTGTATTTTTTTTTTTATTTATATTTATTTTATTATATTTTTTTTTTTTTTTATTTTT 1002

Qy        735 CTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTT 794
                ||  ||||| | ||    ||| |       |   |||| |   |||||  ||| |   ||
Db       1003 ATTGTATTTTTTTTTTTTTATTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTA 1062

Qy        795 ATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTAT 854
              ||   || ||||  ||| | | | | ||     |||||  ||        |    | || |
Db       1063 TTATTTATTTTTTTTTTTTTTTTTATATTTTTTTTTTTTTTTTTTTTTTTTTATTATTTTT 1122

Qy        855 GATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTC 914
                ||  |            |   | |   || |   | ||  |  |   || | || ||
Db       1123 TTTTTTATATTTTTTTTATTTTATTTTTTATTTTATTATTTGTTATTTATTTTTTTTTAT 1182

Qy        915 CTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATT 974
```



RESULT 15
BZ696936
LOCUS BZ696936 1074 bp DNA linear GSS 02-JUL-2003
DEFINITION SP__Ba0091G12.f SP__Ba Sorghum propinquum genomic clone SP__Ba0091G12 5', genomic survey sequence.
ACCESSION BZ696936
VERSION BZ696936.1 GI:28391438
KEYWORDS GSS.
SOURCE Sorghum propinquum
ORGANISM Sorghum propinquum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Wing,R., Yu,Y., Kim,H.R., Collura,K., Pries,G., Currie,J., Soderlund,C. and Hatfield,J.
TITLE Sequencing of Sorghum BAC ends.
http://genome.arizona.edu/stc/sorghum
JOURNAL Unpublished (2003)
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers
FORWARD: atc agc ggc cgc gat cc
BACKWARD: gta aaa cga cgg cca gtg
Plate: 0091 row: G column: 12
Seq primer: atc agc ggc cgc gat cc
Class: BAC ends.
FEATURES Location/Qualifiers
source 1. .1074
/organism="Sorghum propinquum"
/mol_type="genomic DNA"
/db_xref="taxon:132711"
/clone="SP__Ba0091G12"
/clone_lib="SP__Ba"
/note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII; Paterson lab BAC library (HindIII)"
ORIGIN

Query Match 7.4%; Score 149; DB 8; Length 1074;
Best Local Similarity 46.9%; Pred. No. 3.7e-11;
Matches 466; Conservative 0; Mismatches 520; Indels 7; Gaps 1;



```
Db        715 TTTTTTTTTTTTTATTTTTTTTTTTTATTTTTATTTTTTTTTTTTTTTTTTTATTTTTTT 774

Qy       1286 AGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTA 1345

Db        775 TTTTTTTTTTTATCTTTTTTTTTATTTTTTTTTTATATTTATTTTTTATTTATTTTTTTT 834

Qy       1346 TATTCGAATAAATCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAA 1405

Db        835 TTTTTTTATTTTTTTTTTTTTTTTATTAATTTTTTTTTTTTTTTTTTTTTATTATATTTAA 894

Qy       1406 TATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATA 1465

Db        895 TTTTTTTTCATTTTTTTTTTTTTATTTATTTTTTTTTTTTTTTTATTTTTTTATTATTTTT 954

Qy       1466 TGATAAAAATAAATTTTAAATATTTTTATATTT 1498

Db        955 TTATTATAATATATATTTTTTTTTATTATATTT 987
```

Search completed: September 3, 2005, 12:52:46
Job time : 4934.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 1, 2005, 13:28:41 ; Search time 247.78 Seconds
(without alignments)
13286.738 Million cell updates/sec

Title: US-10-776-889-1
Perfect score: 2012
Sequence: 1 atcttaggcccttgattata.....aactaagaagcaatacttca 2012

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	155.6	7.7	18773	4	US-09-949-016-14164	Sequence 14164, A
c	2	135.4	6.7	187169	4	US-09-949-016-12776	Sequence 12776, A
c	3	135.4	6.7	191569	4	US-09-949-016-15940	Sequence 15940, A
	4	134	6.7	18773	4	US-09-949-016-14164	Sequence 14164, A
c	5	122.4	6.1	612	4	US-09-902-540-1357	Sequence 1357, Ap
c	6	122.4	6.1	95255	4	US-09-949-016-17067	Sequence 17067, A
c	7	114.2	5.7	119153	4	US-09-949-016-12378	Sequence 12378, A
	8	110	5.5	29717	4	US-09-949-016-16284	Sequence 16284, A
	9	110	5.5	60376	4	US-09-949-016-12423	Sequence 12423, A
	10	109.4	5.4	30820	4	US-09-949-016-17145	Sequence 17145, A
	11	105.8	5.3	147382	4	US-09-949-016-14624	Sequence 14624, A
c	12	105.6	5.2	30820	4	US-09-949-016-17145	Sequence 17145, A
	13	105.4	5.2	134987	4	US-09-949-016-15348	Sequence 15348, A
	14	105.4	5.2	134987	4	US-09-949-016-15349	Sequence 15349, A
	15	105.4	5.2	134987	4	US-09-949-016-15350	Sequence 15350, A
	16	105.4	5.2	134987	4	US-09-949-016-15507	Sequence 15507, A
	17	105.4	5.2	134987	4	US-09-949-016-15508	Sequence 15508, A
	18	105.4	5.2	134987	4	US-09-949-016-15509	Sequence 15509, A
c	19	104.8	5.2	55886	4	US-09-949-016-15129	Sequence 15129, A
c	20	104	5.2	1039	4	US-09-902-540-1280	Sequence 1280, Ap
	21	103.6	5.1	95255	4	US-09-949-016-17067	Sequence 17067, A
	22	103	5.1	119153	4	US-09-949-016-12378	Sequence 12378, A
	23	102.6	5.1	298336	4	US-09-949-016-16600	Sequence 16600, A
c	24	102.4	5.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
	25	102.2	5.1	47781	4	US-09-949-016-16492	Sequence 16492, A
	26	102.2	5.1	47781	4	US-09-949-016-16493	Sequence 16493, A
	27	102.2	5.1	47781	4	US-09-949-016-16494	Sequence 16494, A
c	28	102	5.1	640681	4	US-09-790-988-1	Sequence 1, Appli
	29	101.8	5.1	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	30	101.4	5.0	47781	4	US-09-949-016-16492	Sequence 16492, A
c	31	101.4	5.0	47781	4	US-09-949-016-16493	Sequence 16493, A
c	32	101.4	5.0	47781	4	US-09-949-016-16494	Sequence 16494, A
	33	101	5.0	12313	4	US-09-949-016-13248	Sequence 13248, A
c	34	100.6	5.0	114139	4	US-09-949-016-16536	Sequence 16536, A
c	35	99.6	5.0	25067	4	US-09-949-016-11794	Sequence 11794, A
c	36	99.6	5.0	25441	4	US-09-949-016-14232	Sequence 14232, A
c	37	97	4.8	59519	4	US-09-949-016-13504	Sequence 13504, A
c	38	96.8	4.8	19438	4	US-09-949-016-12699	Sequence 12699, A
c	39	96.4	4.8	205044	4	US-09-949-016-15851	Sequence 15851, A
c	40	96.4	4.8	205044	4	US-09-949-016-15852	Sequence 15852, A
c	41	96.4	4.8	205044	4	US-09-949-016-15853	Sequence 15853, A
c	42	96.4	4.8	223471	4	US-09-949-016-12387	Sequence 12387, A
c	43	96.4	4.8	223471	4	US-09-949-016-12724	Sequence 12724, A
c	44	96.4	4.8	223471	4	US-09-949-016-12725	Sequence 12725, A
	45	96.2	4.8	19438	4	US-09-949-016-12699	Sequence 12699, A

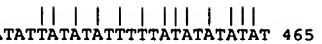
ALIGNMENTS

RESULT 1
US-09-949-016-14164/c
; Sequence 14164, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14164
; LENGTH: 18773
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14164

Query Match 7.7%; Score 155.6; DB 4; Length 18773;
Best Local Similarity 49.5%; Pred. No. 5.9e-18;
Matches 547; Conservative 0; Mismatches 544; Indels 15; Gaps 5;

```
Qy      494 ATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTA 553
            ||| |  ||||||  | | ||| | |     |   |   ||  | | || ||  |   ||
Db    18193 ATAATAATTATATAATGATTTATATAATAATTATATATAAATATATAATTTATATTTATA 18134

Qy      554 TTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATG 613
            | || |   |  |   | |||   | ||  | || | |  || |  ||||       ||
Db    18133 TAATTTATATAATAATTATATATAAATATATATAATATATAAATTATATTATATATCATA 18074

Qy      614 TAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTG 673
            |    || |  |||| |   |     | | |||| | |||  ||| | | | | | | |
Db    18073 TTGTATATAAAATATATAATATTATATATAATATATATGTAATGTATAATATATATAATA 18014

Qy      674 ATATGAAAAATTTATTCTTTTTAAAGC-TATATCATTTTATATATACTTTTTCTTTTCTT 732
              |||  || || |||| | | ||||   ||||  ||| |||||     | |   |  ||
Db    18013 TTATATAATATATATTATATATAAAATATATAATATTATATATAATAGTATATATAAATT 17954

Qy      733 TTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAAT 792
            |  | |   | |  | |  || | |||   |||| | | | |||| | |   | ||  ||
Db    17953 TCATATATATATAAAATATATATCATATATAATATAATATATAAAATATATATCATATAT 17894

Qy      793 TTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATT 852
```

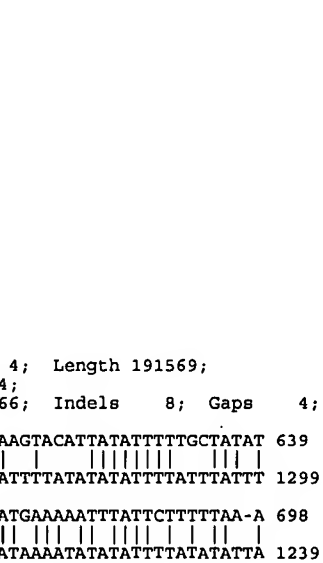
```
Db      17893 AATATAATATATATTATATATTATATATAATATTATATATTATATATATTATATATTATA 17834

Qy        853 ATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATT 912

Db      17833 TATAATATAATATATATTATATATTATATATTATATATTATATATTATATATTATATATT 17774

Qy        913 TCCTAAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCA 972

Db      17773 ACATATAATATATATATAATATATAATATATATTATATATTATACATTATATATTATATA 17714

Qy        973 TTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTA 1032

Db      17713 TTATACATTATATATTATATTATATATTATACATTTATATATTATATTATATATTATATA 17654

Qy       1033 CTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTCCAAAAT 1092

Db      17653 TTATATTATATATTTTTATATTATATATTATATTATATATTGTATTATATATTATATATT 17594

Qy       1093 TGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAACTGAAA 1152

Db      17593 AT----ATTATATATTATATACTATATTTTATATAATATAATATATAGAATACATATTAT 17538

Qy       1153 AGCATTGGCAAGTCTCCTCCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCATCTGATT 1212

Db      17537 ATTATATATAA-TATATAATATATAAAATATATTATATTATATATAAAATATATATTATA 17479

Qy       1213 GATTCAGTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAACTATTAATTAAT 1272

Db      17478 TATTAATTATATATAATTGAT------AATTATGTATAAGATATATTATATATTAATTAT 17425

Qy       1273 CTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAATATGTTATTTT 1332

Db      17424 ATATAATTAATAATTATGTATAAAATATATATTATATATTAATTATATAATTAATAATTA 17365

Qy       1333 TTGTTTAGATGTATATTCGAATAAATCTAAATATATGATA---ATGATTTTTTATATTGA 1389

Db      17364 TATATAAAATATATATTATATTATATATAAAAATATTATATATAAAATATTATATATTAA 17305

Qy       1390 TTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACACATAATTTT 1449

Db      17304 ATAATATTATTTTATATGTTATATATTATACATTATTAGATATTCTATATTATATTTTAT 17245

Qy       1450 ATAAGGATAAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACGAGAAAAAA 1509

Db      17244 ATAATTATATATAATTATATACACATTTTATATAGAAAATATTATATATAATATTATATA 17185

Qy       1510 AAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATTCATAAATT 1569

Db      17184 TAATATTTTATATATAAAATTATATATAAAATATTATATATAATATTATATATAAAATAT 17125

Qy       1570 CCTATATGTATATTTGAAATTAAAAA 1595

Db      17124 TATATATATATAGAAAATAAAGAAAA 17099


RESULT 2
US-09-949-016-12776/c
; Sequence 12776, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12776
;   LENGTH: 187169
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(187169)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12776

  Query Match             6.7%;  Score 135.4;  DB 4;  Length 187169;
  Best Local Similarity   49.4%;  Pred. No. 2.3e-14;
  Matches  463;  Conservative    0;  Mismatches  466;  Indels    8;  Gaps    4;

Qy        580 TAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATAT 639

Db       1358 TATATATATTTTTTATATATATATTTTATATATTTATTTTATATATATTTTATTTATTT 1299

Qy        640 ACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAA-A 698

Db       1298 ATATATTTATATATTATATATTTTTATATATAATATATAAAATATATATTTTATATATTA 1239

Qy        699 GCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAAT 758

Db       1238 TATATAAATATATATATTATATTTATATATATATTTATATATTATATATAATATATAT 1179

Qy        759 AAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGT 818

Db       1178 AATATATATGTATTATATATATTTATATATATAATATATTATATATTATATATTATAT 1119

Qy        819 TTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAG 878

Db       1118 ATTATATATTTTATATATTATATATTATATATTATATATT---ATATATTTTATACATTA 1062

Qy        879 TGAGTAATAAGTGATGAGAAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATG 938

Db       1061 TATATTATATATTATATATTTTATATATTATATATTTTATATTTTATATTTTATATATAT 1002

Qy        939 TATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATT 998

Db       1001 TATATATTTTATATATCATATATATTATATATCCATATATTTTTTATATATTATATATTT 942

Qy        999 TTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCC 1058

Db        941 TATATATTATATATTATATATATTATATATTTTATATATATTATATATTTTATATATTTT 882

Qy       1059 TCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTA 1118

Db        881 ATATATTTTATATATATTTTATATATTTTATATATTATATATATTTTATATATTATATAT 822

Qy       1119 TCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCT 1178

Db        821 TTTATATATTTTATATATTTTATATATTATATATAATATATTTTATATTTTATATATA 762

Qy       1179 CAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTA 1238

Db        761 TTATATATTTTATATTTTTTATATATATTATATATTTTATATTTTTATATATATTATATA 702

Qy       1239 AAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATT 1298

Db        701 TTTATATATATTTTATGTATATTATATATTATATATATTTTATATATTATATATTATATA 642

Qy       1299 TTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAAT 1358

Db        641 TATTTTATATATTATATATTATATATATTTATATATATTATATAGTATATATATTTTATAT 582

Qy       1359 CTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGAT 1418

Db        581 ATA---CTATATAATATATATTTTATATATATTATATATTATATATATTTTATATATTAT 525

Qy       1419 AT-TTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAA 1477
```



```
Db        524 ATATATTTTATATATATTTTTTATATATTTTATATATATTATATATTTTTATATATATAT 465

Qy       1478 ATTTTAAATATTTTTATATTTACGAGAAAAAAAAATA 1514
              || || ||||| |     ||  |||| | ||| | |
Db        464 TTTATATATATTATATATATTCAGAGAGAGAAAGAGA 428
```

RESULT 3
US-09-949-016-15940/c
; Sequence 15940, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15940
; LENGTH: 191569
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(191569)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15940

Query Match 6.7%; Score 135.4; DB 4; Length 191569;
Best Local Similarity 49.4%; Pred. No. 2.3e-14;
Matches 463; Conservative 0; Mismatches 466; Indels 8; Gaps 4;

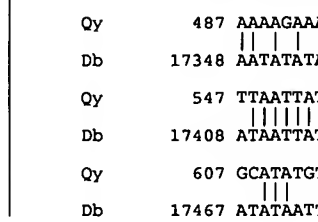


```
Qy        580 TAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATAT 639
Db       1358 TATATATATTTTTTATATATATATTTTATATATTTTATTTTATATATATTTTATTTATTT 1299

Qy        640 ACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAA-A 698
Db       1298 ATATATTTATATATTATATATTTTTATATATAATATATAAAATATATATTTTATATATTA 1239

Qy        699 GCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTAAT 758
Db       1238 TATATAAATATATATATATTATATTTATATATATATTTATATATTATATATAATATATAT 1179

Qy        759 AAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGT 818
Db       1178 AATATATATGTATTATATATATTTATATATATAATATATTATATATTATATATATTATAT 1119

Qy        819 TTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAG 878
Db       1118 ATTATATATATTTATATATTATATATTATATATTATATATT---ATATATTTTATACATTA 1062

Qy        879 TGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATG 938
Db       1061 TATATTATATATTATATATTTTATATATTATATATTTTATATTTATATTTTATATATAT 1002

Qy        939 TATCTACTCTCTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATT 998
Db       1001 TATATATTTTATATATCATATATATTATATATCCATATATTTTTTATATATTATATATTT 942

Qy        999 TTTTTATCATATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCC 1058
Db        941 TATATATTATATATTATATATATTATATATTTTATATATATTATATATTTTATATATTTT 882
```

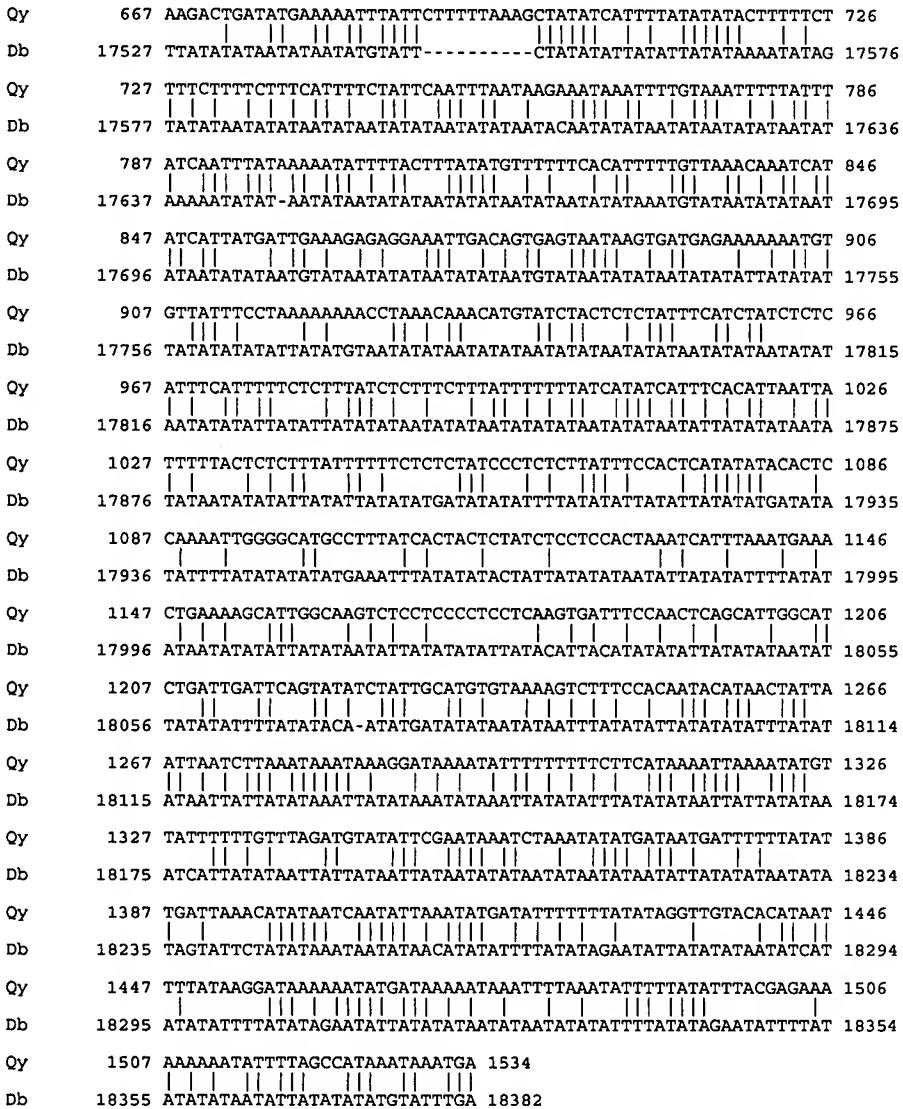
```
Qy       1059 TCTCTTATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCTA 1118
Db        881 ATATATTTTATATATATTTTATATATTTTATATATTATATATATTTTATATATTATATAT 822

Qy       1119 TCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCT 1178
Db        821 TTTATATATTTTATATATTTTATATATTATATATAATATATATTTTATATTTTATATATA 762

Qy       1179 CAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGTA 1238
Db        761 TTATATATTTTATATTTTTTATATATATTATATATTTTATATTTTTATATATATTATATA 702

Qy       1239 AAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATATT 1298
Db        701 TTTATATATATTTTATGTATATTATATATTATATATATTTTATATATTATATATTATATA 642

Qy       1299 TTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAAT 1358
Db        641 TATTTTATATATTATATATTATATATATTTATATATATTATATAGTATATATATTTATAT 582

Qy       1359 CTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGAT 1418
Db        581 ATA---CTATATAATATATATTTTATATATATTATATATTATATATATTTTATATATTAT 525

Qy       1419 AT-TTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAA 1477
Db        524 ATATATTTTATATATATTTTTTATATATTTTATATATATTATATATTTTTATATATATAT 465

Qy       1478 ATTTTAAATATTTTTATATTTACGAGAAAAAAAAATA 1514
Db        464 TTTATATATATTATATATATTCAGAGAGAGAAAGAGA 428
```

RESULT 4
US-09-949-016-14164
; Sequence 14164, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14164
; LENGTH: 18773
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14164

Query Match 6.7%; Score 134; DB 4; Length 18773;
Best Local Similarity 48.2%; Pred. No. 3e-14;
Matches 505; Conservative 0; Mismatches 530; Indels 13; Gaps 4;

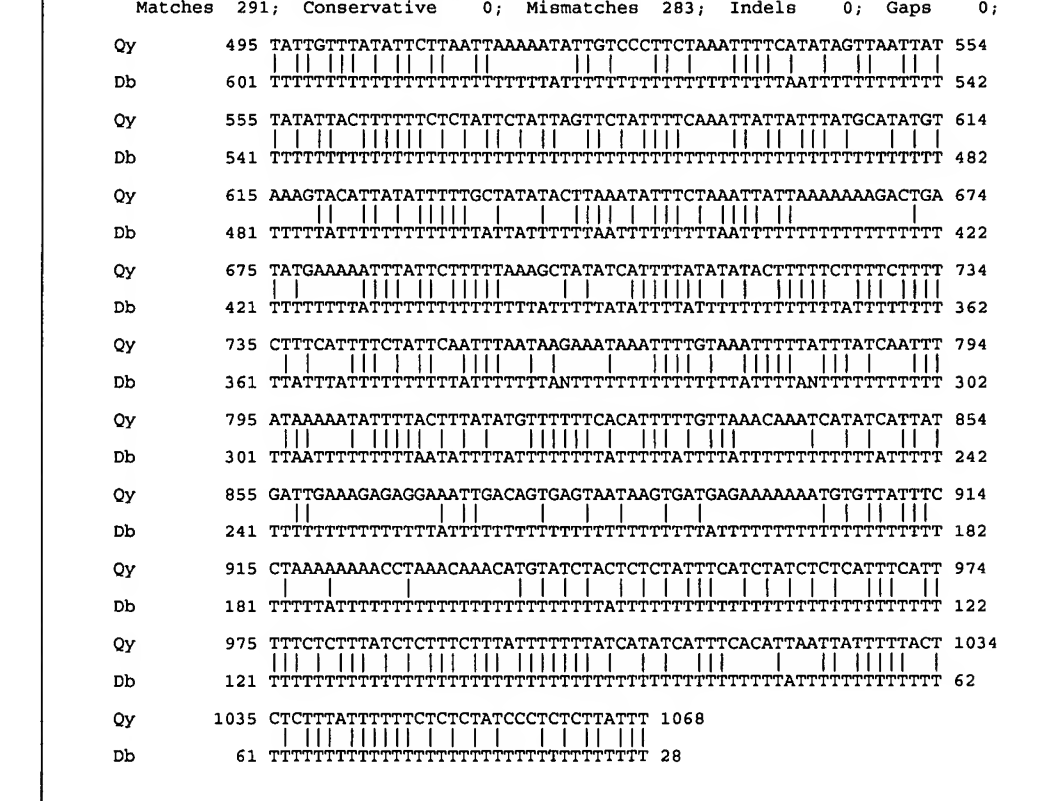


```
Qy        487 AAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAG 546
Db      17348 AATATATATTTTATATATAATTATTAATTATATAATTAATATATAATATATATTTTATAC 17407

Qy        547 TTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTAT 606
Db      17408 ATAATTATTA-ATTATATATAATTAATATATAATATATCTTATACATAATTATCAATTAT 17466

Qy        607 GCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAA 666
Db      17467 ATATAATTAATATATAATATATATTTTATATATAATATAATATATTTTATATATTATATA 17526
```



RESULT 5
US-09-902-540-1357/c
; Sequence 1357, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1357
; LENGTH: 612
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(612)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1357

Query Match 6.1%; Score 122.4; DB 4; Length 612;
Best Local Similarity 50.7%; Pred. No. 1.9e-12;
Matches 291; Conservative 0; Mismatches 283; Indels 0; Gaps 0;



RESULT 6
US-09-949-016-17067/c
; Sequence 17067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17067
;   LENGTH: 95255
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(95255)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17067

  Query Match             6.1%;  Score 122.4;  DB 4;  Length 95255;
  Best Local Similarity   46.1%;  Pred. No. 3.6e-12;
  Matches  520;  Conservative    0;  Mismatches  601;  Indels    7;  Gaps    3;

Qy        488 AAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGT 547
              | | || | |  | |||||| |  ||  || |||   | |  |   | |||  ||||| |
Db      43253 ATACAATTCTGTTATATATTTTGTATATAATATACAATTCTGTTATATTTTGTATATAAT 43194

Qy        548 TAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATG 607
                |  ||| | ||| | |||| | |||  |||     |   ||  |  || |      ||
Db      43193 ATACAATTCTGTTA-TATTTTGTATATAATATATAATTCTGTTATATTTTGTATACAATA 43135

Qy        608 CATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAA 667
              |||  |    || ||||||||||   ||||   |  |||  || ||  || |  | ||
Db      43134 TATAATTCTGTTATATTATATTTTGTATATAATATATAATTGTTATATTTTGTATATAAT 43075

Qy        668 AGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTT 727
              | |    ||   |  || ||||  | | |   |||| | | ||||     |  |  ||
Db      43074 ATATAATTACATATTATATATTAATAATTATTATATATTAATATATAATATATAATAATT 43015

Qy        728 TTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTA 787
              |  |   |    | |  ||  |  |||  | | || |||  |  |||| |    | |
Db      43014 ATTAATATATAATATATTAATATATAATAAATATATATTAATAATTAAAATATAATATAA 42955

Qy        788 TCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATA 847
              |    |  || |  |  | ||   ||||   |  |  |    | |        | || ||
Db      42954 TTCTATATTATATTTTGTATATAATATAAAATCATATATTATATATAATTCTATATTATG 42895

Qy        848 TCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTG 907
              | ||  |   |  |  | | |   |||       | |   ||||  |    |  | |  ||
Db      42894 TTATATTTTGTATATAATATATAATCATATATAATATATAATTCTATATTATATTTTGTA 42835

Qy        908 T-TATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTC 966
              | || |   |||  | |   || | | |  | ||| |||| | ||| | || |||    ||
Db      42834 TATAATATATAATTATATATTATATATAATTCTATATACTTTGTATATAATATATAATTC 42775

Qy        967 ATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTA 1026
              ||   |||| | | ||         ||   || | |  ||| |  | |   ||  | |
Db      42774 TGTTATATTTTGTATATAATATACAATTCTGTTATATTACATTTTGTATATAATATACAA 42715

Qy       1027 TTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACACTC 1086
              || |        ||| | | |  | |   | | |||| |    |      |     | |
Db      42714 TTCTGTTATATTTTGTATATAATATACAATTCAATGTTATATTTTGTATATACCATTGTT 42655

Qy       1087 CAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGAAA 1146
                | |||        | ||    |  | | |    |  | || || |||     | |
Db      42654 ATATTTTGTATATAATATATAATTGTTATATTTTGTATATAATATATAATTCTGTTATAT 42595

Qy       1147 CTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGCAT 1206
                |        |   |  | |  | ||       |    ||||    |   |  ||      |
Db      42594 TATATTTTGTATATAATATATAATTCCTGTTATATTATATTTTGTATATAATATATAATT 42535

Qy       1207 CTGATTGATTCA-----GTATATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAAC 1261
              |||  |  |||       ||||||   |    | | | | | | |  | |||| ||||
Db      42534 CTGTTATATTATATTTTGTATATAATTCTGTTATATTATATTTTGTATATAATATATAAT 42475

Qy       1262 TATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAA 1321
              ||| | |  ||  || || |||||    || |  || | |  |   | ||| |     | |
Db      42474 TATAATTATATTATATATTAATAATTATTATATAATATATAATATATAATATATAATATA 42415

Qy       1322 TATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTT 1381
              ||   ||| |     |  ||| ||| |      | || | |      | ||| | || |
Db      42414 TAATATATATAAATATATGATATATTTATATTAATATATGATATATTTATATTAATATAC 42355

Qy       1382 TATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACAC 1441
                ||||  |||| |   |||    ||||  | |||  || | | | |||||  |      |
Db      42354 GATATATATTATATTAATATATGATATATATTATATTAATATATCATATATTTATATAAT 42295

Qy       1442 ATAATTTTATAAGGATAAAAAATATGATAAAAATAAATTTTAAATATTTTTATATTTACG 1501
              |||  ||     |  ||| || ||||  || | | | || ||| |||  | |||| |
Db      42294 ATATATTATATAAAATATAATATATATTATATAAATATATTATATAAATATATAATATAT 42235

Qy       1502 AGAAAAAAAAATATTTTAGCCATAAATAAATGACCAGCATATTTTACAACCTTAGTAATT 1561
              |  | | |||    ||       |||  |    |  |    ||| || |   |     |||
Db      42234 ATTATATAAATATATTATATATTAATAATTAAAATATATAATTCTATATTATATTATATT 42175

Qy       1562 CATAAATTCCTATATGTATATTTGAAATTAAAAACAGATAATCGTTAA 1609
                |  | |||||   | | |   ||       |  ||||   |||
Db      42174 TTGTATCTAATATATAATTCTATATTATATTTTGTATCTAATATATAA 42127


RESULT 7
US-09-949-016-12378/c
; Sequence 12378, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12378
;   LENGTH: 119153
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(119153)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12378

  Query Match             5.7%;  Score 114.2;  DB 4;  Length 119153;
  Best Local Similarity   48.4%;  Pred. No. 9.6e-11;
  Matches  474;  Conservative    0;  Mismatches  488;  Indels   17;  Gaps    5;

Qy        530 TCTAAATTTTCATATAGTTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATT 589
              | || ||||| || |||   |  | ||| |   | ||||  | |   | |||  | || |
Db      87184 TATATATTTTGATCTAGAACAAAAGTATCTGGGTGTTTTAGCAAAATTTTTAAATATACT 87125

Qy        590 TTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATAT 649
              |  | |||||    |||  || | | || || | ||||| |||    |||||  ||  |
Db      87124 TAAATATTATATACTATATATTTATGAA-TAAACTATATATTTATAATATAAATATGTTA 87066

Qy        650 TTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATT 709
              |   | || |||| || |     |    | | | |     ||| || |  || ||
Db      87065 TATCATATAATTATAATATATTAAAATATATAATATAAATTTATATATTTACATTTAC 87006

Qy        710 TTATATATACTTTTTCTTTTCTTTTCTTTCAT-TTTCTATTCAATTTAATAAGAAATAAA 768
              |||||  ||  |    | |  |  |  |  || ||| ||    |  || ||||  ||| |
Db      87005 TTATAATTATGTAACATATATTAATAATATATATTTATAAATATATTTATAAATAATTAT 86946
```

```
Qy        769 TTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACAT 828
              |  ||| ||  |   | ||  ||| ||| | |   | |   ||||||  || |    ||
Db      86945 ATAAGTATATAATATATAATGTATTAATATATAATATATTAATTATATACTTGTATTAAT 86886

Qy        829 TTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAA 888
              | |  || |  ||| ||||  ||||   | | | | |   |  |   |  ||      |
Db      86885 ATATAATATATTAATTATATACTTATATATAATATATACTTATATATAACAGATACTATA 86826

Qy        889 GTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCT 948
              | |   | | ||| | | || |    | ||    || | ||  || ||
Db      86825 TATAAGTATATATAATATATAATATATCTTATAATTAATATATAATAATATAATTAAA 86766

Qy        949 CTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCAT 1008
              ||  |     | | || |  ||| | | ||||| | | | ||   | | ||   |
Db      86765 TATATTATTTATATTTTTATATATTTTATATATTTATATATATTTTATATATATTTATAA 86706

Qy       1009 ATCATTTCACATTAATTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCT----- 1063
              || |    | || |  ||| ||   | |  |   || |  | | ||| | | | |
Db      86705 ATAAAAATATATAATATATATTATATTTTATGTATTATATTTTATATACTTATATAATAT 86646

Qy       1064 ------TATTTCCACTCATATATACACTCCAAAATTGGGGCATGCCTTTATCACTACTCT 1117
                    ||||   | |  ||||   |  || ||      |   ||    | || |
Db      86645 AATATATATTATATATTAATTATAGTATATAATATATTACATTATATTACATATTATTAC 86586

Qy       1118 ATCTCCTCCACTAAATCATTTAAATGAAACTGAAAAGCATTGGCAAGTCTCCTCCCCTCC 1177
              | | |   | | ||   || |  |  | |     | | | |     |
Db      86585 TATGTAATATATTACTATGTAATATATTACATAGTAATAATATGTAATATAATGTATTAA 86526

Qy       1178 TCAAGTGATTTCCAACTCAGCATTGGCATCTGATTGATTCAGTATATCTATTGCATGTGT 1237
              | | ||   ||  |  | |    ||  || | ||  ||    || || ||||  || | |
Db      86525 TAATGTAGATTATATATTATATATGATATATAATATATAATATAGATATATTATATAT-T 86467

Qy       1238 AAAAGTCTTTCCACAATACATAACTATTAATTAATCTTAAATAAATAAAGGATAAAATAT 1297
              | |  |  ||  | |   ||   |  ||  |||   || |   |||  |  ||| |   |
Db      86466 ATATATACTTATATATAGTATCTGTTATATATAAGTATATAAGTATATATTATATATATT 86407

Qy       1298 TTTTTTTTCTTCATAAAATTAAAATATGTTATTTTTTGTTTAGATGTATATTCGAATAAA 1357
              | | |  ||  || | ||  |     ||| | | | |  || ||||| |  || |
Db      86406 AATATATATTTATTACATATATTAATATATATAATATATATTAATATATATACTTATTAT 86347

Qy       1358 TCTAAATATATGATAATGATTTTTTATATTGATTAAACATATAATCAATATTAAATATGA 1417
              ||||||  ||||  || | |||||||    || || ||||||  ||||  | |   |
Db      86346 ATATTATATATTATAAGTATATATTATATTATATATACTTATAATATATATACACTTATA 86287

Qy       1418 TATTTTTTTATATAGGTTGTACACATAATTTTATAAGGATAAAAAATATGATAAAAATAA 1477
              ||   | ||||||   | ||  ||| || |||||   |||| | |||| | | | ||||
Db      86286 TA---TAATATATAAAATATAATCATTATATTATATTTATAACATATATAAAATATATAA 86230

Qy       1478 ATTTTAAATATTTTTATAT 1496
              | |  |  ||   |||||
Db      86229 GTATATATAATAAATATAT 86211


RESULT 8
US-09-949-016-16284
; Sequence 16284, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16284
;   LENGTH: 29717
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(29717)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16284

  Query Match             5.5%;  Score 110;  DB 4;  Length 29717;
  Best Local Similarity   53.3%;  Pred. No. 4.3e-10;
  Matches  324;  Conservative    0;  Mismatches  275;  Indels    9;  Gaps    4;

Qy        493 AATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATT 552
              | |  | | |||||| |||  ||   ||||      || || ||||  ||||| |||  |
Db      15340 ATTTATATATATATTTTTATATATTTATATATATATTTATATATTTATATATATTTATAT 15399

Qy        553 ATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATTGCATAT 612
              ||  |   ||  ||| | | ||||    |   | |||||  | ||  ||||||||  ||||
Db      15400 ATATTTATATATTTATATATATTTATATTTATTTATTTATATATATTTATTTATATATAT 15459

Qy        613 GTAAAG-TACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGAC 671
              |  |  || |||    | |   | ||||  ||  ||| | || |||  || | |   |
Db      15460 ATTTATATATATTTATATATATATTTATATATATTTATATATATATTTATATATATTTAT 15519

Qy        672 TGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCT 731
              ||||  | |  |   | ||| || |  |||||   ||||||||||| | ||| | | |
Db      15520 ATATATTTATATGTATATTTTTATATATTTATATATATTTATATATACATATTCATATAT 15579

Qy        732 TTTCTTTCAT-TTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCA 790
              || ||  || ||| ||| ||  || |||   | |    | | |  || | ||  |||
Db      15580 ATTTTTATATATTTATATACATATTCATATATATTTTTATATATTTATATATACATATTT 15639

Qy        791 ATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTT----TTGTTAAACAAATCAT 846
              || |||| | ||||||   | |||| | |||  | || |     | || |   |  |||
Db      15640 ATATATATACATATTTATATATATACATATTTATATATATACACATATTTATATATACAT 15699

Qy        847 ATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGT 906
              ||  ||||  ||  |        | ||| | |       |||  |      | | | || |
Db      15700 ATATTTATATTTATATATATTTATATTTATATATTTATATATTTATATTTATATATATTT 15759

Qy        907 GTTATTTCCTAAA---AAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCT 963
              |  ||   || |  | | |  || | | |  | ||| ||     |  |  || | | |
Db      15760 ATATTTATATATTTATATATTTATATATATTTATATTTATATATTTATATATATTTAT 15819

Qy        964 CTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAA 1023
              |  ||  |||| | | ||||  | | | | ||| ||| |||      ||||    |||
Db      15820 ATATTTATATTTATATATTTACATATATTTATATATTTATATTTATATATTTATATTTAT 15879

Qy       1024 TTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACA 1083
              ||||| ||      ||   ||| | | |   |  | |  | |   || | |||| | |
Db      15880 ATATTTATATATATTTATATTTATATATAATTGATTTTTGTCTACACAATGATATTTTTA 15939

Qy       1084 CTCCAAAA 1091
                || |||
Db      15940 TACCTAAA 15947


RESULT 9
US-09-949-016-12423
; Sequence 12423, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12423
; LENGTH: 60376
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(60376)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12423

Query Match 5.5%; Score 110; DB 4; Length 60376;
Best Local Similarity 53.3%; Pred. No. 4.6e-10;
Matches 324; Conservative 0; Mismatches 275; Indels 9; Gaps 4;

```
Qy        493 AATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATT 552
              | |  | | |||||| |||  ||   ||||       || || ||||  ||||| |||  |
Db      47379 ATTTATATATATATTTTTATATATTTATATATATTTATATATTTATATATATTTATAT 47438

Qy        553 ATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATAT 612
              ||   |  ||  ||| | | ||||    |   | |||||  | ||  |||||||  ||||
Db      47439 ATATTTATATATTTATATATATTTATATTTATTTATTTATATATATTTATTTATATATAT 47498

Qy        613 GTAAAG-TACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGAC 671
              |  |  || |||    | |   | ||||  ||  ||| | || |||  || | |   |
Db      47499 ATTTATATATATTTATATATATATTTATATATATTTATATATATATTTATATATATTTAT 47558

Qy        672 TGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCT 731
              ||||  | |  |   | ||| || |  |||||   ||||||||||| | |||  | | |
Db      47559 ATATATTTATATGTATATTTTTATATATTTATATATATTTATATATACATATTCATATAT 47618

Qy        732 TTTCTTTCAT-TTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTTATTTATCA 790
              || ||  || ||| ||| ||  || |||   | |    | | |  || | ||  |||
Db      47619 ATTTTTATATATTTATATACATATTCATATATATTTTTATATATTTATATATACATATTT 47678

Qy        791 ATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTT----TTGTTAAACAAATCAT 846
              || |||| | ||||||    | ||||  | |||  | || |      | ||  |   | |||
Db      47679 ATATATATACATATTTATATATATACATATTTATATATATACACATATTTATATATACAT 47738

Qy        847 ATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGT 906
              ||  ||||  ||  |     | |||  | |      |||  |      | | | || |
Db      47739 ATATTTATATTTATATATATTTATATTTATATATTTATATATTTATATTTATATATATTT 47798

Qy        907 GTTATTTCCTAAA---AAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCT 963
              |  ||   || |   | | |  || | | |  | ||| ||      |  |  || | | |
Db      47799 ATATTTATATATATTTATATATTTATATATATTTATATTTATATATTTATATATATTTAT 47858

Qy        964 CTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAA 1023
              |  ||  |||| | | |||| | | | | ||| ||| |||      ||||    |||
Db      47859 ATATTTATATTTATATATTTACATATATTTATATATTTATATTTATATATTTATATTTAT 47918

Qy       1024 TTATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACA 1083
              ||||| ||     ||  ||| | | |   | |  | |   || | |||| | |  |
Db      47919 ATATTTATATATATTTATATTTATATATAATTGATTTTTGTCTACACAATGATATTTTTA 47978

Qy       1084 CTCCAAAA 1091
              || |||
Db      47979 TACCTAAA 47986

RESULT 10
US-09-949-016-17145
; Sequence 17145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17145
; LENGTH: 30820
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17145

Query Match 5.4%; Score 109.4; DB 4; Length 30820;
Best Local Similarity 47.0%; Pred. No. 5.4e-10;
Matches 470; Conservative 0; Mismatches 521; Indels 8; Gaps 4;

```
Qy   487 AAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAG 546
              |||| | ||| | | || || || | | ||||| | | || | |||||
Db   5863 AAAATATATAATATATAAATATATATTATATAATATATAATATATAAATATATTATAT 5922

Qy   547 TTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTAT 606
              | || || | | | | || | ||| | | || | || || | |||
Db   5923 ATTATATATAATATTATATAATATATAATATATAAATATATATTATATATAATATTATAT 5982

Qy   607 GCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAA 666
              | | | ||||| || |||| ||| ||| | ||| |||| |
Db   5983 AATATATAATATATAAATATATATTATATATAATATTATATAATATATAATATATAAATA 6042

Qy   667 AAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCT 726
              | | ||| | || ||| | | |||| ||||| || | |||||| | | |
Db   6043 TATAT---TATATACTATATATAATATATAAATATATATTATATACTATATATAATATAT 6099

Qy   727 TTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTT 786
              | | || || | ||| ||| || || | ||| | | | | | ||| |
Db   6100 AAATATATATTATATACTATATATAATATATAAATATATATTATATACTATATATAATAT 6159

Qy   787 ATCAATTTATA-AAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCA 845
              || ||| |||| | ||| | || | |||| | | | | | | || | | |
Db   6160 ATAAATATATATTATATACTATATATAATATATAATATATACTATATTATATATAGTATA 6219

Qy   846 TATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATG 905
              | || | | | | || || | | | || | || | || | |||
Db   6220 TTATATACTATATTATATAGTATAATATAGAATATATTATATATTGTATATAAGATAATA 6279

Qy   906 TGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCAT-CTATCTC 964
              | |||| | | | | || || | | | ||| | | | || || ||| |
Db   6280 TATTATATATTGTATATAAGATATAATATATTATATATTGTATATAAGATATAATATATT 6339

Qy   965 TCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAAT 1024
              || | | | | | || | || | | | | || |||| || | || ||
Db   6340 ATATATTGTATATAAGATAATATATTATATATTGTATATAAGATATAATATATTATATAT 6399

Qy   1025 TATTTTTACTCTCTTTATTTTTTCTCTCTATCCCTCTCTTATTTCCACTCATATATACAC 1084
               | | | || | | | || | | | ||| | ||||||
Db   6400 TGTATATAAGATATAATATATTATATATTGTATATAAGATATAATATATTATATATTGTA 6459

Qy   1085 TCCAAAATTGGGGCATGCCTTTATCACTACTCTATCTCCTCCACTAAATCATTTAAATGA 1144
               | || || | ||| | | | ||| | |
Db   6460 TATAAGATATAATATATTATATATTGTATATAAGATATAATATATTATATATTGTATATA 6519

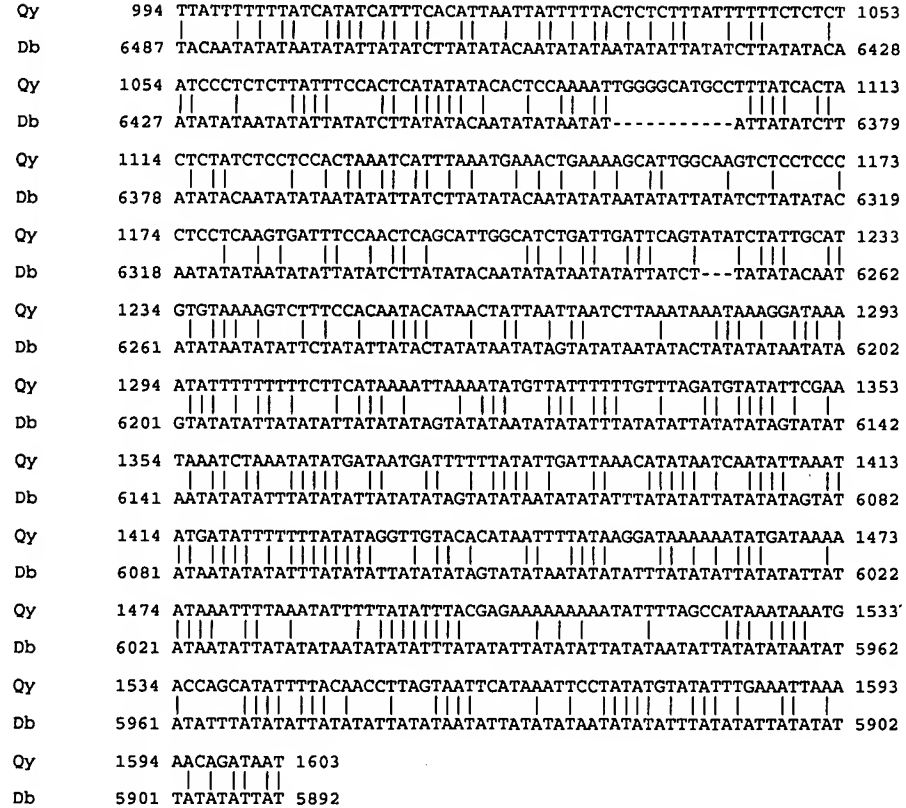
Qy   1145 AACTGAAAAGCATTGGCAAGTCTCCTCCCCTCCTCAAGTGATTTCCAACTCAGCATTGGC 1204
               | | || ||| | | | | | | | | | | |
Db   6520 AGATATAATATATTATATATTGTATATAAGATATAATATATTATATATTGTATATAAGAT 6579

Qy   1205 ATCTGATTGATTCAGTAT---ATCTATTGCATGTGTAAAAGTCTTTCCACAATACATAAC 1261
               || | || ||| ||| || || | || | || | | ||| |
Db   6580 ATATTATATATTGTATATAAGATATAATATATTATATATTGTATATAAGATATAATATAT 6639

Qy   1262 TATTAATTAATCTTAAATAAATAAAGGATAAAATATTTTTTTTTCTTCATAAAATTAAAA 1321
               ||| || || |||||| || | | | ||||| || | | ||| | || | |
```
```

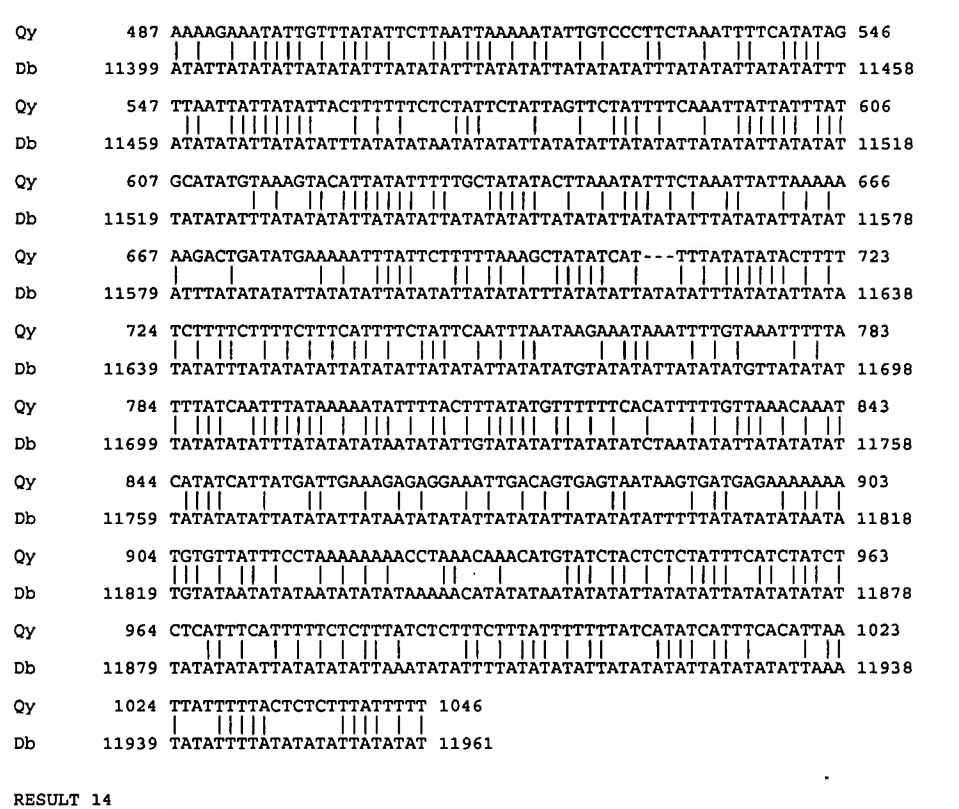
```
Db       6640 TATATATATATTATAAATATGTATATTTTTATATATTATTATAAATATATATATTTTATA 6699

Qy       1322 TATGTTATTTTTTGTTTAGATGTATATTCGAATAAATCTAAATATATGATAATGATTTTT 1381
              ||| ||||| |   | || || | |||    || | | |||||||||  | ||   | ||
Db       6700 TATATTATTATAAATATATATTTTTATATATATTATTATAAATATATATTTATATATATT 6759

Qy       1382 TATATTGATTAAACATATAATCAATATTAAATATGATATTTTTTTATATAGGTTGTACAC 1441
              |||||  |          | || | ||    ||   | || |  | |      |  |
Db       6760 TATATATAAATTTACATGTGTTAAAAATATGATTGTGTTCTTGTATTCTTACAATTGTAT 6819

Qy       1442 ATAATTTTATAAGGATAAAAAATATGATAAAAATAAATT 1480
              |   ||          |  |||||||  ||||||| | |
Db       6820 ACGTTTCCCATTTCCCACTTAATATGAAAAAAATACACT 6858


RESULT 11
US-09-949-016-14624
; Sequence 14624, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14624
;   LENGTH: 147382
;   TYPE: DNA
;   ORGANISM: Human
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(147382)
;   OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14624

  Query Match           5.3%;  Score 105.8;  DB 4;  Length 147382;
  Best Local Similarity   52.5%;  Pred. No. 2.7e-09;
  Matches  299;  Conservative    0;  Mismatches  267;  Indels    3;  Gaps    3;

Qy        497 TTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAGTTAATTATTA 556
              || |||||||   |   |  | ||||  |   || || || || || ||  || |||||
Db     136698 TTATTTATATATATTTATTTATATATATTTATTTATATATATTTATTTATATATTTATTT 136757

Qy        557 TATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAA 616
                |  ||||| | ||||   |||  | |||||  | ||  |||||||||  || | |
Db     136758 ATATTTATTTTTATATATTTATTTATATATATTTGTATATATTTATTTATATATTTATTT 136817

Qy        617 AGTACATTATAT-TTTTGCTATATACTTAAATATTTCTAAATTATTAAAAAAAGACTGAT 675
              |    |||| | | |   ||||||  | | |   | || |||  || |   | |   ||
Db     136818 ATATATTTATTTATATATTTATATATATTAGTTCATATATATTTATATATTTATATATT 136877

Qy        676 ATGAAAAATTTATTCTTTTTAAAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTC 735
              |  | || ||||  | |   || |||||   |||||||||| ||||| | |  || |
Db     136878 TAGTTATATATATTTATATATTAGTTATATATATTTATATATA-TTTTTATATATTTATT 136936

Qy        736 TTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTA 795
              | |   | | |||  | |   |||   |   |  ||  || |||| ||| ||   || ||
Db     136937 TATATATATGTATATATTAACATATTTATATATATTCATATATTTATATATAAATATATA 136996

Qy        796 TAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAATCATATCATTATG 855
              ||  ||||| ||  | |||| | | |    || | |||| |   | | ||| |||  |
Db     136997 TATTAATATATTTATATATATTTATATATTAATATATGTTTATATATTTATA-CATATTT 137055

Qy        856 ATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCC 915
```

```
              ||  | |   |   |  |      |   || | |  |||     | |     |||| |
Db     137056 ATGTATATTAATATATATATTTATATATATTTATATATGTATTTATATATATTTATATAT 137115

Qy        916 TAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCTCTCATTTCATTT 975
              | | |   |  || | |   || ||| |  |   |   | ||| | |   || || |
Db     137116 TTATATTTATTTATATATGTATTTATACATATTTATATATTTATATTTATTTATTTATAT 137175

Qy        976 TTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAATTATTTTTACTC 1035
               | | | ||| | | | ||||||| |||||       | |      |  ||| ||||
Db     137176 ATATTTATATATTTATATTTATTTATTTATATATATTTATATATATTTATATATTTATAT 137235

Qy       1036 TCTTTATTTTTTCTCTCTATCCCTCTCTT 1064
              |  ||  ||| | | | | |   | | ||
Db     137236 TTATTTATTTATATATATTTATATATATT 137264


RESULT 12
US-09-949-016-17145/c
; Sequence 17145, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17145
;   LENGTH: 30820
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-17145

  Query Match           5.2%;  Score 105.6;  DB 4;  Length 30820;
  Best Local Similarity   46.9%;  Pred. No. 2.4e-09;
  Matches  483;  Conservative    0;  Mismatches  524;  Indels   23;  Gaps    4;

Qy        577 TATTAGTTCTATTTTCAAATTATTATTTATGCATATGTAAAGTACATTATATTTTTGCTA 636
              ||  | |   || || || || |   | ||       | | |  |   ||| |||||  ||
Db       6901 TAAAAATAACATCTTAAAGTTGTCTCTAAGAACATTACATATGAGTGTATTTTTTTCATA 6842

Qy        637 TATACTTAAATATTTCTAAATTATTAAAAAAAGACTGATATGAAAAATTTATTCTTTTTA 696
              |  | |   | ||    ||  |||  ||      |          || |       ||||
Db       6841 TTAAGTGGGAAATGGGAAACGTATACAATTGTAAGAATACAAGAACACAATCATATTTTT 6782

Qy        697 AAGCTATATCATTTTATATATACTTTTTCTTTTCTTTTCTTTCATTTTCTATTCAATTTA 756
              ||   || | | ||||||||||  | |    |  | |   || ||  |      ||| ||
Db       6781 AACACATGTAAATTTATATATAAATATATATAAATATATATTTATAAT------AATATA 6728

Qy        757 ATAAGAAATAAATTTTGTAAATTTTTATTTATCAATTTATAAAAATATTTTACTTTATAT 816
               | ||||| || ||   |||  |   |    ||| |||| |  |||  || | ||||
Db       6727 TATAAAAATATATATTTATAATAATATATATAAAATATATATATTTATAATAATATATAA 6668

Qy        817 GTTTTTTCACATTTTTGTTAAACAAATCATATCATTATGATTGAAAGAGAGGAAATTGAC 876
              | | || |||| |  || | | || ||||  |      ||  |    |   |    |
Db       6667 AAATATACATATTTATAATATATATATAATATATTATATCTTATATACAATATATAATAT 6608

Qy        877 AGTGAGTAATAAGTGATGAGAAAAAAATGTGTTATTTCCTAAAAAAAACCTAAACAAACA 936
              | |    |  || |          | || || | |||| | | | | | ||  ||    | |
Db       6607 ATTATATCTTATATACAATATATAATATATCTTATATACAATATATAATATATTATATCT 6548

Qy        937 TGTATCTACTCT---CTATTTCATCTATCTCTCATTTCATTTTTCTCTTTATCTCTTTCT 993
              | |||  | | |    ||| | || | |      |   || |      |||| ||||
Db       6547 TATATACAATATATAATATATTATATCTTATATACAATATATAATATATTATATCTTATA 6488
```



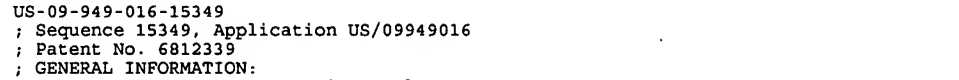
RESULT 13
US-09-949-016-15348
; Sequence 15348, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15348
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15348

Query Match 5.2%; Score 105.4; DB 4; Length 134987;
Best Local Similarity 50.4%; Pred. No. 3.2e-09;
Matches 284; Conservative 0; Mismatches 276; Indels 3; Gaps 1;



RESULT 14
US-09-949-016-15349
; Sequence 15349, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15349
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15349

Query Match 5.2%; Score 105.4; DB 4; Length 134987;
Best Local Similarity 50.4%; Pred. No. 3.2e-09;
Matches 284; Conservative 0; Mismatches 276; Indels 3; Gaps 1;



```
Db      11399 ATATTATATATTATATATTTATATATTTATATATTATATATATTTATATATTATATATTT 11458

Qy        547 TTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTAT 606
              ||  ||||||||   | | |    |||    |   |  ||| |   |  |||||| |||
Db      11459 ATATATATTATATATTTATATATAATATATATTATATATTATATATTATATATTATATAT 11518

Qy        607 GCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAA 666
              | |  || |||||||  ||    |||||  |  | ||| | |  ||   | | |
Db      11519 TATATATTTATATATATTATATATTATATATATTATATATTATATATTTATATATTATAT 11578

Qy        667 AAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCAT---TTTATATATACTTTT 723
              |    |       | |  ||||   || || |  |||||  |   | | ||||||  | |
Db      11579 ATTTATATATATTATATATTATATATTATATATTTATATATTATATATTTATATATTATA 11638

Qy        724 TCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTA 783
              | | ||  | | | | || |  |||  | | ||     | |||   | | |    | |
Db      11639 TATATTTATATATATTATATATTATATATTATATATGTATATATTATATATGTTATATAT 11698

Qy        784 TTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAAT 843
              | |||  ||||||| | ||| | || | ||||| || | |   |  | | ||| | | ||
Db      11699 TATATATATTTATATATATAATATATTGTATATATTATATATCTAATATATTATATATAT 11758

Qy        844 CATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAA 903
              ||||   |   ||   | | |   | |  | | |   ||     | ||    | || |
Db      11759 TATATATATTATATATTATAATATATATTATATATTATATATATTTTTATATATATAATA 11818

Qy        904 TGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCT 963
              ||| | || |   | | | |    ||   |     ||| || | | ||||  || ||| |
Db      11819 TGTATAATATATAATATATATAAAAACATATATAATATATATTATATATTATATATATAT 11878

Qy        964 CTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAA 1023
              || |  | | | | || |     || | ||| | ||   |||| || |    | ||
Db      11879 TATATATATTATATATATTAAATATATTTTATATATATTATATATATTATATATATTAAA 11938

Qy       1024 TTATTTTTACTCTCTTTATTTTT 1046
              |   |||||      |||| | |
Db      11939 TATATTTTATATATATTATATAT 11961
```

RESULT 15
US-09-949-016-15350
; Sequence 15350, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15350
; LENGTH: 134987
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15350

Query Match 5.2%; Score 105.4; DB 4; Length 134987;
Best Local Similarity 50.4%; Pred. No. 3.2e-09;
Matches 284; Conservative 0; Mismatches 276; Indels 3; Gaps 1;

```
Qy        487 AAAAGAAATATTGTTTATATTCTTAATTAAAAATATTGTCCCTTCTAAATTTTCATATAG 546
              | |  | ||||| | ||| |   || ||| | ||  | |   ||   |  ||  ||||
Db      11399 ATATTATATATTATATATTTATATATTTATATATTATATATATTTATATATTATATATTT 11458

Qy        547 TTAATTATTATATTACTTTTTTCTCTATTCTATTAGTTCTATTTTCAAATTATTATTTAT 606
              ||  ||||||||   | | |    |||    |   |  ||| |   |  |||||| |||
Db      11459 ATATATATTATATATTTATATATAATATATATTATATATTATATATTATATATTATATAT 11518

Qy        607 GCATATGTAAAGTACATTATATTTTTGCTATATACTTAAATATTTCTAAATTATTAAAAA 666
              | |  || |||||||  ||    |||||  |  | ||| | |  ||   | | |
Db      11519 TATATATTTATATATATTATATATTATATATATTATATATTATATATTTATATATTATAT 11578

Qy        667 AAGACTGATATGAAAAATTTATTCTTTTTAAAGCTATATCAT---TTTATATATACTTTT 723
              |    |       | |  ||||   || || |  |||||  |   | | ||||||  | |
Db      11579 ATTTATATATATTATATATTATATATTATATATTTATATATTATATATTTATATATTATA 11638

Qy        724 TCTTTTCTTTTCTTTCATTTTCTATTCAATTTAATAAGAAATAAATTTTGTAAATTTTTA 783
              | | ||  | | | | || |  |||  | | ||     | |||   | | |    | |
Db      11639 TATATTTATATATATTATATATTATATATTATATATGTATATATTATATATGTTATATAT 11698

Qy        784 TTTATCAATTTATAAAAATATTTTACTTTATATGTTTTTTCACATTTTTGTTAAACAAAT 843
              | |||  ||||||| | ||| | || | ||||| || | |   |  | | ||| | | ||
Db      11699 TATATATATTTATATATATAATATATTGTATATATTATATATCTAATATATTATATATAT 11758

Qy        844 CATATCATTATGATTGAAAGAGAGGAAATTGACAGTGAGTAATAAGTGATGAGAAAAAAA 903
              ||||   |   ||   | | |   | |  | | |   ||     | ||    | || |
Db      11759 TATATATATTATATATTATAATATATATTATATATTATATATATTTTTATATATATAATA 11818

Qy        904 TGTGTTATTTCCTAAAAAAAACCTAAACAAACATGTATCTACTCTCTATTTCATCTATCT 963
              ||| | || |   | | | |    ||   |     ||| || | | ||||  || ||| |
Db      11819 TGTATAATATATAATATATATAAAAACATATATAATATATATTATATATTATATATATAT 11878

Qy        964 CTCATTTCATTTTTCTCTTTATCTCTTTCTTTATTTTTTTATCATATCATTTCACATTAA 1023
              || |  | | | | || |     || | ||| | ||   |||| || |    | ||
Db      11879 TATATATATTATATATATTAAATATATTTTATATATATTATATATATTATATATATTAAA 11938

Qy       1024 TTATTTTTACTCTCTTTATTTTT 1046
              |   |||||      |||| | |
Db      11939 TATATTTTATATATATTATATAT 11961
```

Search completed: September 1, 2005, 22:18:37
Job time : 254.78 secs